JPL PUBLICATION 85-76

Product Assurance Technology for Custom LSI/VLSI Electronics

Report for Period: October 1982 - September 1984

M.G. Buehler
B.R. Blaes
G.A. Jennings
B.T. Moore
R.H. Nixon
C.A. Pina
H.R. Sayah
M.W. Sievers
N.F. Stahlberg

June 1985

Prepared for
National Aeronautics and Space Administration,
Defense Advanced Research Projects Agency,
and
U.S. Department of Defense
by
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

JPL PUBLICATION 85-76

Product Assurance Technology for Custom LSI/VLSI Electronics

Report for Period: October 1982 - September 1984

M.G. Buehler
B.R. Blaes
G.A. Jennings
B.T. Moore
R.H. Nixon
C.A. Pina
H.R. Sayah
M.W. Sievers
N.F. Stahlberg

June 1985

Prepared for
National Aeronautics and Space Administration,
Defense Advanced Research Projects Agency,
and
U.S. Department of Defense
by
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

The research described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, and was sponsored by the Defense Advanced Research Projects Agency, the U.S. Department of Defense, and the National Aeronautics and Space Administration.

ABSTRACT

The work described in this report represents the collaborative efforts of integrated-circuit (IC) parts specialists, device physicists, test-chip engineers, and fault-tolerant-circuit designers. Their efforts were focused on developing the technology for obtaining custom ICs from CMOS-bulk silicon foundries using a universal set of layout rules. In pursuit of this goal the technical efforts were guided by the requirement to develop a 3-μm CMOS test chip for the Combined Release and Radiation Effects Satellite (CRRES). This chip contains both analog and digital circuits for characterizing the space-radiation-induced shifts in CMOS transistor parameters and inverter propagation delays, and for characterizing the single-event-upset (SEU) rates of static random access memories (SRMs). The development employed all the elements required to obtain custom circuits from silicon foundries, including circuit design, foundry interfacing, circuit test, and circuit qualification.

The technical accomplishments of this effort include:

(a) A critical review of the military IC qualification standards. The review indicated that many of these standards are not applicable to the procurement of custom ICs.

(b) An assessment of the test time and area required by CMOS-bulk test structures that are to be included in parameter extraction and yield analysis test chips.

(c) The development of a MOSFET parameter extraction procedure, called JMOSFIT, which allows the extraction of physically meaningful parameters for a SPICE-like circuit simulator.

(d) The fault modeling of pinhole-array-capacitor defects in terms of a gate-to-silicon resistive short that terminates in an n-diffusion pocket in the silicon.

(e) The time response analysis of CMOS gates (inverters, NANDs, NORs, and flip-flops) for various resistive shorts which indicates the resistive values needed to cause a faulty response.

(f) The development of a CMOS-bulk test chip for CRRES.

CONTENTS

Figures

Tables

SECTION 1

INTRODUCTION

1.1 OVERVIEW

The objective of this effort is to develop a product assurance methodology that will allow the procurement of reliable, custom LSI/VLSI circuits from silicon foundries and will permit the use of these circuits in critical system applications such as spacecraft.

The use of test chips in the manufacture of integrated circuits (ICs) is now a routine matter [1, 2]. However, their use in the procurement process has lagged.

The procurement of highly reliable parts by the military was recently discussed by Macaruso [2]. He concluded that the "military procurement technique emphasizes testing and retesting at a time when dramatic improvements in process control obviate the need. In fact it has been shown that extensive testing shortens circuit life by exposing ICs to inappropriate tests as well as to increased risks from electrostatic discharge."

A major conclusion of this effort is that test chips can greatly enhance the procurement of integrated circuits by allowing the identification of wafers that will not meet parametric and defect density specifications before the wafers are scribed and circuits packaged and sent through reliability screening. In addition, test structures can be overstressed so as to identify flaws such as time-dependent dielectric breakdown and electromigration. These potential circuit failure mechanisms cannot be characterized from tests on the circuits themselves.

This effort reviews the usefulness in applying the military standards to the wafer acceptance of custom ICs in Section 2.1, and the results are summarized in Section 2.1.3. Section 2.2 describes a comprehensive test chip for evaluating a CMOS wafer and presents the critical parameters that are derived from the test chip. During the contract period, selected test structures were developed as indicated in Section 2.3. These structures were selected to advance the state of the art in defect detection and yield analysis (pinhole array capacitor), in contact resistor parameter spread analysis (contact resistor process cliff), in inverter parameter spread analysis (addressable inverter matrix), in device reliability (reliability structures), in worst-case circuit design (inverter noise margin analysis), and in MOSFET parameter extraction (JMOSFIT). These structures were fabricated in both 3-μm and 1.2-μm CMOS-bulk, and selected test results are given in Section 2.4. Also discussed in this section are the methods by which the test chips were generated, the test programs were generated, and the data was analyzed. During this period, 24 test chips were developed as indicated in Table 2.4.3-1. In Section 2.5, which is concerned with fault models, we investigated the operating state of open-gate transistors and found that n-channel transistors were in an off state, whereas p-channel transistors were mostly in an on state. In addition, we evaluated the transient response of inverters, NAND gates, NOR gates, flip-flops, and dynamic latches to the presence of

various resistor faults. In Section 2.6 we describe the development of a test chip for the Combined Release and Radiation Effects Satellite (CRRES). This chip is intended to provide device level information on how transistor parameters shift and how integrated circuits behave in the low-dose mixed-particle environment of space. To aid in the analysis, the JPL CRRES chip includes a transistor matrix, a timing sampler circuit, and a 1K or 4K static RAM.

1.2 REFERENCES

1. C. Alcorn, D. Dworak, N. Haddad, W. Hendley, and P. Nixon, "Kerf Test Structure Designs for Process and Device Characterization," Solid State Technol., 28, 229-235 (May 1985).

2. E. Macaruso, "DOD Vexes IC Makers," Electronics Week, 58, 63-68 (February 4, 1985).

3. J. H. Nelson and H. L. Chew, "Test Device for CMOS/SOS Parameter Testing," in Presentations of the IEEE VLSI Workshop on Test Structures (February 1984).

SECTION 2

TECHNICAL ACCOMPLISHMENTS

The results of this product assurance technology effort are presented in six sections, as follows: Section 2.1 discusses the wafer acceptance procedures, Section 2.2 the test chips, Section 2.3 the test structures, Section 2.4 the test chip methodology, Section 2.5 the fault models, and Section 2.6 the CRRES test chip.

2.1 WAFER ACCEPTANCE PROCEDURES[1]

As was stated in the previous report for the period May 1981 - October 1982 [1], the procedures for procuring reliable, custom integrated circuits from silicon foundries may require that current procurement practices be modified, and that the traditional approaches for qualifying integrated electronics for high-reliability applications be supplemented with additional procedures such as those employing test-chip based measurements. The reason on which these assertions are based is the apparent inadequacy of current procurement practices, and traditional qualification approaches, to address recent advances both in integrated circuit technology and in the applications engineering environment. These advances are several, and include the advent of very-large-scale integration (VLSI) and the attendant very-small-scale feature sizes, the perceived need to utilize custom circuitry as well as standard devices, and the emerging separation of integrated circuit design and fabrication, with the latter function being provided by the silicon foundry.

During the current contract period (October 1982 - October 1984), qualification procedures currently employed by different organizations have been investigated to ascertain their status and applicability relative to the scenario described above. Particular attention has been given to those areas in which additional, useful information can be provided by employment of the test-chip methodology developed as part of this program. Development of prototype documentation has begun, to provide formal requirements and procedures that incorporate those features of both the traditional approach and the test-chip methodology that are both applicable and proven. In the overall qualification process, many traditional procedures are both applicable (perhaps with some modification) and proven, while few traditional procedures deal with wafer acceptance from a silicon foundry. As a result, the prototype wafer acceptance requirements and procedures are based almost entirely on the test-chip methodology and are similar to those currently employed by ISI [2], with additional reliability-related factors.

2.1.1 Basis for Evaluation of Current Qualification Procedures

2.1.1.1 Introduction. The evaluation of existing qualification procedures for potential application to custom VLSI electronics requires close examination of several critical issues surrounding the nature of these devices. A basic assertion is that VLSI is both quantitatively and qualitatively different from lesser scales of integration, and procedures that may be well suited for small and medium-scale integration or discrete devices may be totally inappropriate for VLSI. The issue of customization adds an additional dimension for examination.

2.1.1.1.1 Very-Large-Scale Integration (VLSI). The term "very-large-scale integration," or VLSI, infers three interrelated attributes: high circuit complexity, small feature size, and relatively large chip size. The utilization of VLSI in place of small or medium-scale integration can result in a

[1]This section was prepared by Norman F. Stahlberg.

substantial increase in system reliability, due to a drastic decrease in external interconnections; however, such utilization also complicates reliability assurance matters greatly. At present, no one has developed a qualification/screening method demonstrably optimized to assure VLSI device reliability. Some work has been done by Schroen [3] but, as will be seen below, his use of process controls may not be applicable in the foundry environment. Many of the techniques employed in conventional qualification/screening procedures are oriented toward small and medium-scale level technology. VLSI implies a level of complexity that may not allow complete electrical testing of the device (unless adequate testability features are designed into it) and that surely obscures knowledge of many internal states during life test and burn-in. An additional implication at the present time is the possible use of fabrication process technologies that are not fully mature.

2.1.1.1.2 Highly Automated Design and Testing. To deal with the large size and high complexity of VLSI devices in a timely and cost-effective manner, the capability of the individual designer must be enhanced or supplemented through the use of automated design and layout tools, and the respective capability of the test engineer enhanced by automated testing techniques. This inserts a new dimension into the problems of reliability assurance and device qualification, but also presents a new and potentially powerful set of techniques for use in achieving those ends. Current procedures for qualification and reliability assurance do not yet take adequate advantage of this valuable resource.

2.1.1.1.3 High Reliability Products. The integrated circuits and subsystems addressed in this report are for application in systems that must exhibit extreme reliability. The integrated devices contained in those systems must offer long service life (on the order of ten years) and high resistance to hostile environments, including both temperature and ionizing radiation. The issues of accelerated life testing and burn-in need to be revisited within the context of VLSI devices and associated technologies.

2.1.1.1.4 Custom Circuits and Subsystems. The term "custom" implies three interesting and perplexing attributes of the devices being addressed: low-volume production (relative to a typical commercial volume), lack of user history, and "foundry" fabrication. Conventional qualification/screening methods assume large (cumulative) numbers of devices will be constructed such that statistical analysis can be applied. Although some of this analysis is possible given a smaller total population, the validity of these approaches is questionable with small numbers of devices. Implicit in most conventional and proposed methods is a dependence on the time and volume of production so as to permit design improvements and identification of deviant lots by comparison to historical data (i.e., by using a learning curve). The low-volume situation will limit the utility of the learning curve to improve the device as volume accumulates, and gives rise to the use of accelerated life testing to supply the missing basis for reliability prediction. Also missing is the ability to use information from a large body of collateral users to feed back to the design engineers regarding the functionality and performance of the device in field applications. As a result, additional effort must be given to design validation and verification before deployment of the devices (preferably before

fabrication). The issue of foundry fabrication is taken up in the next section.

2.1.1.1.5 Nontraditional Development Methodology. The issue of custom devices leads to another new element not addressed by conventional qualification procedures, namely the contemporary employment of a nontraditional methodology in the development of these devices. This methodology involves the general separation of development functions between different organizations. These functions include logic design, circuit design, circuit layout, lithography, wafer fabrication, device packaging, and testing. In this context the wafer fabrication is performed by a "silicon foundry," and often an additional organization, the "silicon broker," is included to coordinate the different functional activities.

Foundries can vary greatly from those who will provide only wafer fabrication service to those who will fabricate, test, and package your device. The foundry environment assumed in this evaluation will provide more than just minimal service, but will not in any way be a full-service type. It is assumed that this foundry will limit the access to the process by the customer and will permit no external control over the process. Thus any conventional qualification/screening methods that rely explicitly or implicitly on access to or control of the process in any direct way will not be applicable to the environment as described above.

2.1.2 Review and Evaluation of Military Qualification Procedures

Much of the material presented in this section is derived from research conducted at the University of Arizona under the auspices of the JPL Director's Discretionary Fund (JPL Contract Number 956855) [4].

2.1.2.1 Introduction. The practices and procedures reviewed in this report are represented by the following documents [5-11], which are reviewed individually in the sections that follow:

1. MIL-STD-883C
2. MIL-STD-976A
3. MIL-STD-977
4. MIL-M-38510F
5. MIL-HDBK-339 (USAF)
6. MIL-STD-CLSIC (NWSC/ICE) [proposed document]
7. MIL-STD-XXXX (RADC/ITT) [proposed document]

This review has been made within the context set forth in Section 2.1.1, which focuses on the ability of the procedure or practice to satisfy the reliability requirements of custom, high-reliability VLSI devices. The individual document reviews are divided into two parts: first, a description of the purpose and content of the document, and second, a commentary on details identified as being inappropriate, irrelevant, outdated, or deficient in some regard. In summary, given the context set forth in Section 2.1.1, problems were felt to exist at three levels:

1. Basic assumptions. The basic assumptions underlying much of the MIL qualification process as represented by the documents listed above would appear to be flawed. For example, the existence of process controls does not necessarily lead to a reliable product, nor do the controls specified in the current documentation lead to a uniform, repeatable process.

2. Procedural concepts. Because of these alleged flaws in the basic assumptions, many resulting concepts are ineffective (e.g., line certification and production lot acceptance).

3. Technical requirements. Many of the current technical requirements are either inadequate or inapplicable, since many major failure mechanisms and reliability parameters are ignored, and many measurement and inspection techniques are seriously outdated. Technical requirements are changing rapidly, both as feature sizes shrink and as our understanding of various physical phenomena progresses.

A further summary of the detailed evaluations may be found in Section 2.1.3.

2.1.2.2 Document Evaluations.

2.1.2.2.1 MIL-STD-883C: Test Methods and Procedures for Microelectronics.

A. DESCRIPTION

The stated purpose of this document is to "establish uniform methods, controls, and procedures for designing, testing, identifying and certifying microelectronic devices suitable for use within military and aerospace electronic systems." The coverage of the document is rather comprehensive; the sections of the document that are particularly relevant to the subject at hand, and are commented upon herein, are:

1. Method 1010.5: Temperature Cycling
2. Method 1014.6: Seal
3. Method 1015.5: Burn-In Test
4. Method 2010.7: Internal Visual Inspection
5. Method 2012.5: Radiography
6. Method 2018.1: Scanning Electron Microscope (SEM) Inspection of Metallization
7. Method 5007.5: Wafer Lot Acceptance
8. Method 5010: Test Procedures for Custom Monolithic Microcircuits

B. COMMENTS

1. Method 1010.5: Temperature Cycling

The present temperature limits are specified to be -65 degrees C to +150 degrees C. Modern manufacturing technology, however, often uses either vapor phase or infrared reflow of solder for board attachment of integrated circuits. These reflow methods require temperatures of at least 210 degrees C. Thus, the upper temperature limit for the temperature cycling screen test is substantially lower than the temperature to which the circuit will be cycled (at least once, assuming no rework) in the system manufacturing environment.

2. Method 1014.6: Seal

Maintenance of the cavity dew point below 0 degrees C for an extended period of time requires a leak rate well below the fine seal leak rate of 5 X 10 exp-8 atm-cc/sec permitted by this method [12]. In addition, test condition B allows the use of the Krypton method of leak detection. Harmful effects to the die due to radiation have been experienced using this leak detection method in cases where marginal seal leaks are involved [13]. Lowering the fine seal leak rate limit by at least a factor of five should minimize both problems, and is easily supported by current technology.

3. Method 1015.5: Burn-In Test

Burn-in is the most effective screen to eliminate infant mortalities, because it simulates actual or worst-case operation of the device accelerated through a time/power/temperature relationship. It is most effective for detecting die-related failure modes, which vary with the process technology used in the manufacture of the device.

In complex devices such as VLSI microcircuits, there is little distinction between the stresses resulting from static or dynamic burn-in, since it becomes increasingly difficult to implement a clear-cut version of either option. The complexity of VLSI tends to isolate and mask the internal states of the chip. The static burn-in configuration tends to stress only the input and output circuitry, without providing the intended stress to interior circuit nodes. By the same token, reverse-bias burn-in becomes essentially meaningless, due to the impossibility of knowing which input state reverse-biases the maximum number of junctions.

The best knowledge to date indicates that some combination of static and dynamic burn-in, combined with an overvoltage stress test for MOS devices, yields the best burn-in screen. The overvoltage test is not a true burn-in technique, per se, but is more effective in detecting oxide breakdown (a major failure mode in MOS devices) than is temperature stressing. This is due to the low activation energy, and resulting weak thermal dependence, of the oxide breakdown mechanism.

The burn-in specifications given in Method 1015.5 need to be examined closely and modified in light of the reasons mentioned above. Additionally, the general issue of burn-in must be placed on firmer footing, taking into account the failure modes and mechanisms introduced by increasingly smaller feature sizes and new technologies. Ways of analyzing new devices to determine the proper burn-in parameters must be devised, since the burn-in condition developed for a particular device must be based on an understanding of that device's construction, operation, and topography [14].

4. Method 2010.7: Internal Visual Inspection

The utility of this test was greatest when SSI and MSI technology was the state of the art. Given the feature sizes and element densities typical of VLSI, the practicality of this test is limited to gross defects on the die or in bonding, or contamination in the package.

5. Method 2012.5: Radiography

The use of x-ray techniques can be useful in determining defects in the die-attach material for an encapsulated device. The value of this test is highly dependent on the experience of the operator evaluating the resulting photographs. Therefore, an improperly trained or skilled operator can provide inadequate screening for die-attach defects. In addition, this test is not useful for evaluating bond-wire integrity unless the bond wires are exclusively gold [15]. Finally, this test can be used for seal-width measurements, but again, the adequacy of the test is operator-dependent. The utility of this screen is thus questionable.

6. Method 2018.1: Scanning Electron Microscope (SEM) Inspection of Metallization

The intent of this method is to assure the minimum current-carrying capability of the chip metallization. This is important to detect potential failure mechanisms due to electromigration. The method depends on the inference of wafer lot metallization attributes from semi-quantitative observations by SEM on a small sample of lot metallization. The generally subjective interpretation of the observations is known and is one drawback of the technique. A more serious drawback is the fact that pass/fail criteria are not directly linked to long-term performance of the device, but are usually based on worst-case assumptions regarding electromigration mean-time-to-failure (MTTF) and metal thinning at steps. It is possible to implement test structure-based methods that more directly and quantitatively predict future performance of the metallization. This more

direct measurement method would allow some line widths that appear marginal from SEM inspection to pass and, more importantly, reject metal coverages that look acceptable in SEM views, but in fact are not capable of handling the given current.

7. Method 5007.5: Wafer Lot Acceptance

The wafer acceptance phase represents a crucial step in foundry-fabricated custom device development. Contrary to the traditional scenario, the foundry has no responsibility for the design, only for its implementation into silicon. In general, the customer must have control of the acceptance process and, hence, needs additional insight into, if not responsibility for, the wafer-level testing. The following tests need to be examined for applicability to the custom microcircuit scenario and appropriateness of limits to contemporary fabrication technology.

a. Wafer Thickness - According to the precept that too thin a wafer will lead to brittleness and too thick a wafer will make scribing difficult to accomplish without introducing defects, this method is valid. From the standpoint of microcircuit reliability, however, this test is of little consequence.

b. Metallization Thickness - The intent of this test is to determine the current-carrying capability of the metal. What the test will not detect is the uniformity of the thickness (e.g., step coverage), variations in which can have deleterious effects on metal performance regardless of thickness. Refer to comments in item (6) above (concerning Method 2018.1).

c. Thermal Stability - The weakness of this specification lies in the retention of criteria that do not reflect the capabilities of present semiconductor technology. For example, the limit for gate oxide stability in MOS devices is specified to be 400 mV. A limit in the realm of 50 mV to 100 mV is more suitable for high-reliability devices and is well within the capabilities of modern, well-controlled processes utilizing 0.1-micrometer gate oxide thicknesses.

d. Gold Backing Thickness - Measurement of this parameter is appropriate only if eutectic die attach methods are employed. Currently, the use of noneutectic adhesives is being considered as a means of solving the problems encountered in eutectically attaching the large die used in VLSI devices [16].

8. Method 5010: Test Procedures for Custom Monolithic Microcircuits

This method replaces 5004 (Screening Procedures) and 5005 (Qualification and Quality Conformance Procedures) for the custom microcircuit scenario. As with 5004, it refers back to 5007.5 for coverage of Wafer Lot Acceptance (see comments above).

What Method 5010 adds for application to custom microcircuits is a provision for die evaluation (including radiation testing per Methods 1017.2, 1019.2 and 1020), die attachment evaluation, and (optional) pre-seal burn-in. Method 5010 removes reverse-bias burn-in as a requirement, but retains it as an option. It also increases the burn-in time from 240 hours to 320 hours for Class S devices, dividing that time into two 160-hour increments with an intermediate electrical screen. Without doubt, these changes are both appropriate and significant; the retention of reverse-bias burn-in, however, even as an option, is without sound basis. Life testing per Method 1005.4 (1000 hours at 125 degrees C) is retained in Method 5010. A weakness in this procedure is that any failure mechanism not activated within the specified parameters will not be identified. Also, the sampling method employed ensures only that the particular devices tested meet the specified requirements, and says nothing for the other devices in the production run. In addition, life testing does not act as a predictor of reliability; it can only indicate that a failure has occurred.

2.1.2.2.2 MIL-STD-976A: Certification Requirements for JAN Microcircuits.

A. DESCRIPTION

The purpose of this standard is to establish criteria for certification, as required by MIL-M-38510F. This certification is a prerequisite to microcircuit qualification, and is performed in advance of, and independent of, procurement. More specifically, the standard establishes the minimum requirements for the certification of manufacturing facilities and lines used in the production and testing of high-reliability JAN-grade microcircuits. Coverage includes plant facilities, equipment, personnel training, process controls, testing, and documentation. Actual test methods and procedures used to implement these requirements are provided by MIL-STD-977 (refer to Section 2.1.2.2.3 below).

B. COMMENTS

The underlying assumptions followed in 976A are that access exists to the process, that the process can be "frozen," and that modification of the process is possible. Although no "typical" foundry exists, it is in general true that:

1. Details of the process used by the foundry will not be revealed to the customer.
2. The process will not be "frozen" for the customer.
3. The foundry will not expend the time or the money required to obtain line certification.

Because of this (understandable) situation, the requirements embodied in MIL-STD-976A are probably inapplicable to the custom microcircuit scenario.

There are also other objections to the basic tenets of 976A, the most basic of which is that 976A infers that the capability to control critical processes within certain bounds will automatically ensure reliability. No guarantee exists that if a process is "under control," it will produce a reliable product. In addition, the phrase "established bounds" used in 976A is very nebulous; the established bounds for a reliable device are not defined in 976A, and generally are unknown (except perhaps through experience with the particular process in reference to a particular type of device). An additional, and serious, deficiency in 976A is that the standard does not directly address the major failure mechanisms that predominate in VLSI technology. For example, Section 5.3.10 contains specifications for inspection by a scanning electron microscope. The current-carrying capability of metallization of a given chip is only indirectly addressed by this method. Maintaining a SEM program for measuring metallization thickness at oxide steps does not directly ensure that problems at those steps will not occur, since the sample sizes involved in the method are of necessity quite small. The test-structure-based method mentioned above in relation to MIL-STD-883C, Method 2018.1, would use a significantly larger sample size.

2.1.2.2.3 MIL-STD-977: Test Methods and Procedures for Microcircuit Line Certification.

A. DESCRIPTION

This standard describes test methods and procedures applicable to the control of materials and processes used in the manufacture of microcircuits, and is explicitly limited to silicon wafer fabrication. The standard is intended to complement MIL-STD-883, and covers the operations required during wafer processing and inspection, starting with the raw materials and ending with the finished wafer. Some alternative test methods and procedures are also given as references, and the choice from among the listed alternatives is at the manufacturer's option unless otherwise specified in the procurement documentation. Additionally, at the manufacturer's option, equivalent test methods and procedures may be used, provided that the results are within the desired accuracy of measurement and that approval has been granted by the certifying authority.

B. COMMENTS

Although 977 is entitled "Test Methods and Procedures for Microcircuit Line Certification," application of the test methods and procedures contained in it does not seem to be restricted to line certification; indeed, 977 seems to be treated as an extension of MIL-STD-883C, providing test methods and procedures for use in the wafer fabrication phase of JAN microcircuit production. Interestingly, 977 is not explicitly referenced in MIL-M-38510F, as one would expect, but is referred to only indirectly through 883C, where the primary reference is in Method 5007.5 (Wafer Lot Acceptance). Those sections of MIL-STD-977 that are referenced in MIL-STD-883C, Method 5007.5, are as follows:

1. Method 1580: Substrate Thickness and Flatness
2. Method 2500: Capacitance-Voltage Measurements to Determine Oxide Quality
3. Method 5500: Measurement of Film Thickness

2.1.2.2.4 MIL-M-38510F: General Specification for Military Microcircuits.

A. DESCRIPTION

The purpose of this specification is to establish the general requirements for military (JAN) microcircuits, including the quality assurance and reliability assurance requirements that must be met in the acquisition of those microcircuits. The scope of 38510F includes multichip microcircuits and hybrid devices as well as monolithic microcircuits. (The distinction between multichip microcircuits and hybrids is rather subtle; the former is defined as "a microcircuit consisting of elements formed on or within two or more semiconductor chips which are separately attached to a substrate," while the latter is defined as consisting of elements that are a combination of the film microcircuit type and the semiconductor type, or a combination of one or both with discrete parts.)

Major concepts that are embodied in 38510F, and which will be discussed in this report, are as follows:

1. Line Certification
2. Product Assurance Plan
3. Wafer Lot Acceptance
4. Package Seal
5. Screening Tests
6. Device Qualification

B. COMMENTS

Concerns with 38510F for purposes of application to the custom, foundry-fabricated microcircuit scenario can be categorized according to the topics listed in item (A) above. These concerns are as follows:

1. Line Certification

MIL-M-38510F requires line certification as prescribed in MIL-STD-976A. The problems associated with this concept have been discussed in relation to that document. In addition to the unlikelihood of a foundry accepting or actually maintaining a process freeze following certification, the benefits of doing so when the life-of-type is only a few wafer lots is not clear.

2. Product Assurance Plan

In the foundry environment, the requirement of a formal Product Assurance (PA) plan as detailed in 38510F is of questionable feasibility and benefit. Some degree of product assurance surely will be performed on the part of the foundry, but almost certainly not to the extent prescribed in 38510F. Government approval of a quality control plan probably would be unacceptable to a foundry; also, the existence of a PA plan does not ensure that a qualified reliable product will be produced. The existence and implementation of the plan only infers that there will be a resultant increase in the quality of the product being produced. The arguments for the maintenance of such a PA plan, therefore, are similar to those presented in relation to MIL-STD-976A.

3. Wafer Lot Acceptance

According to MIL-M-38510F, a wafer lot is defined as "...consist[ing] of microcircuit wafers processed in a manner that requires every wafer to be subjected to each batch process as a group...." Implicit in this definition is the assertion that uniformity exists between wafers after being subjected to each batch process. Conversely, investigations by Linholm [17] and Buehler [1] have indicated that, in reality, significant intralot variations exist. This is consistent with the fact that few, if any, contemporary fabrication processes are truly batch processes. These results lead to the conclusion that wafer acceptance must be performed on an individual wafer basis; indeed, in some cases, parametric variations across a wafer may be sufficient to require acceptance to be performed on a chip-by-chip basis. In either case, traceability of a die to a wafer lot is meaningless.

Testing methods associated with the wafer lot acceptance provisions of 38510F are given in MIL-STD-883C, Method 5007.5, and individual comments are to be found in reference to that Method (Section 2.1.2.2.1 B-8 of this report). An additional point to be made is that, since large quantities of custom devices would not be produced, the idea of lot jeopardy that is embedded in the wafer lot concept is not applicable. The conclusion that follows this analysis is that wafer lot acceptance as defined in MIL-M-38510F is technically inadequate to meet the requirements of custom microcircuits.

4. Package Seal

Present procedures call for a package sealing environment of 5000 ppm at +100 degrees C maximum moisture content. Noting that 5000 ppm moisture gives a dew point slightly below 0 degrees C, it is widely agreed that the limit should be reduced. Present technology easily supports a limit of 3000 ppm, and 1000 ppm maximum is not difficult to achieve reproducibly at a competent packaging facility.

5. Screening Tests

Given the silicon foundry scenario, the organization responsible for performing the screening tests is not generally prescribed. Once a foundry fabricates wafers associated with a custom device, the processed wafers are generally delivered to the customer, who, in turn, determines the testing and packaging of the die, unless the foundry also offers those services. In the the case of MIL-M-38510F, all screens are assumed to have been performed under the same roof as was the fabrication. Therefore, the direct application of MIL-M-38510F to the foundry situation is unclear.

The screening sequence utilized in 38510F is based on MIL-STD-883C, Method 5010. Many wafer-level tests need to be strengthened or added, including, for example, determination of oxide integrity and time-dependent dielectric breakdown (TDDB) characteristics. Overvoltage stress testing should be added to the burn-in procedure; detailed comments are given in Section 2.1.2.2.1 of this report (concerning MIL-STD-883C, Method 1015.5).

6. Device Qualification

Qualification and quality conformance per 38510F is primarily based on 1000-hour life testing of the product devices as specified in MIL-STD-883C, Method 5010. Implicit in this practice is that if a failure mechanism is not activated within the parameters of these particular tests, the mechanism will not be

identified. Given this problem, and those detailed in Section 2.1.2.2.1 of this report (concerning MIL-STD-883C, Method 5010), life testing may be only marginally beneficial to the qualification program required for custom microcircuits.

As a whole, the applicability of MIL-M-38510F to custom microcircuit qualification is severely limited. No mention is made in 38510F of developing quality and reliability in the design stage of a device. Also, new and more direct methods for determination of quality and reliability of the physical implementation need to be be developed, relative to those specified in 38510F, for application to foundry-fabricated custom microcircuits.

2.1.2.2.5 MIL-HDBK-339 (USAF): Custom Large Scale Integrated Circuit Development and Acquisition for Space Vehicles.

A. DESCRIPTION

This document contains requirements for the management, design, and manufacturing control of custom, monolithic, large-scale, integrated microcircuits intended for use in high-reliability space systems. Additionally, 339 contains general requirements for the use of those devices in such high-reliability applications. As a military handbook, 339 is intended to be referenced as a guidance document in space vehicle acquisition contracts.

The MIL-HDBK is divided into four parts. There is a main section, which does the conceptual development for the program plan, and three appendixes. Appendix C is of particular importance, since it provides the guidelines by which a custom microcircuit would be screened and qualified for space system use. The main body of 339 details designer and manufacturer requirements to ensure that the issue of testability is addressed during the design of the custom microcircuit. The testability program is outlined completely in Appendix A, which describes in detail the philosophy behind the testability issue and the methodology required for successful implementation of a testability program. In addition, Appendix A includes specific design considerations.

Topics covered by 339 which have been selected for comment in this report are as follows:

1. Product Assurance Program
2. Designer Capability Audit
3. Manufacturer Capability Audit
4. Hardness Assurance Program
5. Appendix C: General Military Specification for Large Scale Integrated Circuits for Space Vehicles

B. COMMENTS

For evaluating the reliability of microcircuits, the MIL-HDBK provides a very comprehensive alternative to MIL-STD-38510F. MIL-HDBK introduces some reliability assurance concepts that, to date, have not been presented in other procedures (either accepted or proposed). In the Functional Design section, the MIL-HDBK defines the test structures that would be desirable for usage on test chips. Included are test structures that would be valuable for reliability prediction.

A side-by-side comparison of MIL-HDBK-339 with MIL-M-38510F reveals many similarities between the two, with minor variations in limits and procedures. Some of the notable differences between the two are described in item (5) below. The variations between 38510F and 339 reveal an attempt by the latter to implement methods to better test VLSI devices. Notable is the use of test structures in the MIL-HDBK, even though no specifics are given regarding testing environments. Also, radiation hardness is emphasized in this document.

1. Product Assurance

The product assurance program described in 339 affects both the designer and manufacturer of integrated circuits. The program requires documentation of all requirements associated with the related product assurance. The program also requires manufacturers to continually compare the production device to the one designed conceptually via destructive physical testing on a sample basis. Other details include identifying key inspection points in the manufacturing operation and those manufacturing steps that are critical to reliability assurance.

From a reliability assurance standpoint, product assurance does not assure device reliability. Such a program will ensure that a design and process are being implemented faithfully by the manufacturer, but reliability is assured only through inferential logic. Furthermore, in the custom VLSI environment, the product assurance concept is not applicable. The ability to specify the programs that a foundry will implement to assure that a functioning, reliable device is being manufactured probably will not exist.

2. Designer Capability Audit

A unique characteristic of the MIL-HDBK is that reliability-related requirements are placed on the designer as well as the manufacturer. Some key points of this portion of 339 are as follows:

a. A design baseline is assumed to exist.
b. Information is obtained regarding previous design experience of the design organization.
c. The design organization must have computer-aided design (CAD) capability.
d. The designer must have the capability to include testability and fault tolerance in the custom design.
e. The designer must have the capability to design for radiation hardness.

3. Manufacturer Capability Audit

The manufacturer capability audit is the approximate equivalent of line certification per MIL-STD-976A. Two key points from this section are:

a. A manufacturing and assembly baseline is assumed to exist and is also assumed to be presentable to outside concerns.
b. The manufacturer is required to follow the quality assurance provisions, including the screening procedures, detailed in Appendix C. This, in turn, requires the capability to successfully fulfill the requirements regarding test structures, failure analysis, and hardness assurance.

Because of the similarity of this section of MIL-HDBK-339 to the line certification requirements of MIL-STD-976A, some of the problems inherent to 976A also apply to 339. One of these problems is due to the unwillingness of foundries to reveal details about processes to customers. Also, a foundry may or may not have extensive experience in a particular area (e.g., 3-micrometer, poly-gate CMOS technology) [18]; considering the size of many current foundries, a broad range of experience is not probable. Finally, a typical foundry will not be able to satisfy the requirement for hardness assurance capability. It would appear that a capability audit per MIL-HDBK-339 would eliminate most, if not all, of today's foundries from being considered as sources for high-reliability, radiation-hardened, custom microcircuits. At best, it would identify individual foundries as potential sources of custom fabrication in a specific technology. If the provisions of MIL-HDBK-339 are to be followed, a new generation of silicon foundries will have to be brought into existence to service this specialty market; only time will tell whether or not this will occur. On the other hand, there is the possibility of defining procedures that can effectively deal with the limitations of current foundries, and still ensure a high-quality, high-reliability product.

4. Hardness Assurance Program

The MIL-HDBK prescribes a detailed program for hardness assurance. The program includes the formation of a Hardness Assurance (HA) Organization to create and implement an HA plan.

This organization would have members contributing to design reviews and manufacturer audits. The designer of a radiation hardened device is required to provide detailed information on the performance and testing of the device in a radiation environment. As mentioned above in item (3), a foundry is required to have radiation hardness experience. In the Physical Design section of 339, layout rules are specified for HA. The HA program and its associated features can assist in assuring radiation hardness during the design phase, but in the custom microcircuit environment, the requirements during and after manufacture of the device are of primary concern; therefore, features of the HA program are of secondary importance.

Appendix A (Radiation Hardness Requirements) contains detailed information on a radiation-hardened device lot-conformance program. This program defines the radiation hardness level of a device and establishes three categories of devices related to the radiation design margin. Appendix C (General Military Specification) mentions the use of either test devices or chips for determining levels of radiation hardness for a given device. This appendix establishes the use of test structures for radiation-hardness testing, but the structures are not used as predictors for specific failure mechanisms occurring during the irradiation of a device. Although the radiation-hardness level of a lot has been determined by this method, the reason for the degree of hardness remains unknown. Failure mechanism information can be used to determine weaknesses in the fabrication process, and such information should be obtained and so used. Finally, no method for radiation screening is specified in Appendix C.

5. Appendix C: General Military Specification

In addition to the radiation-hardness aspects mentioned above in item (4), the general specification outlines some specific design criteria, including the following:

a. The use of standard cells and circuit partitioning is emphasized.
b. "Reliance on post-assembly screening tests to detect discrepant devices shall be reduced to the extent practicable, by the use of design features and test structures that can be used for in-process controls, inspections, or tests."
c. Testability features are required to be included in the design whenever possible.
d. Limits are specified for conductors, passivations, and environmental stresses, including temperature and pressure.

The concepts provided by Appendix C are fundamental to good design, and introduce some ideas that have not previously been presented in a formal qualification program. These concepts will improve only the quality of design, however, not the reliability of the product produced by a foundry.

Notable differences between the tenets of MIL-HDBK-339 and MIL-M-38510F include the following:

a. Reduced reliance on post-assembly screening. The MIL-HDBK states explicitly, "Reliance on post-assembly screening tests to detect discrepant devices shall be reduced to the extent practicable by the use of design features and test structures that can be used for in-process controls, inspections, or tests." The emphasis is on building, rather than testing, reliability into a device. The use of specific test structures is detailed to complement this de-emphasis of post-assembly screening. These test structures include both process monitors and reliability predictors. In contrast, 38510F relies heavily on post-assembly screening with no use of test structures.

b. The MIL-HDBK requires dynamic as well as static burn-in to be performed as part of the screening requirements; 38510F requires only static burn-in, although through MIL-STD-883C, Method 5010, allows the option of dynamic. Dynamic burn-in should stress a maximum number of gates, which is an important consideration with highly complex VLSI devices. Developing a dynamic burn-in program, however, is not trivial.

c. The MIL-HDBK provides much more comprehensive coverage of the radiation-hardness issue than does MIL-M-38510F, paying extensive attention to radiation-critical materials, processes, and procedures. Design aspects of radiation hardness also are included. Radiation testing is much more severe than in 38510F, and includes testing to ten times the specification level or failure, whichever comes first, for each specified radiation environment.

d. The concept of a wafer lot is maintained in MIL-HDBK-339, but the size of the lot is limited to four wafers; MIL-M-38510F places no limit on wafer lot size. Since the number of wafers processed together is typically 10 or more, limiting the lot size to four wafers is tantamount to reducing the wafer lot to a single wafer, and should be so stated. Refer to Section 2.1.2.2.4 of this report (concerning MIL-M-38510F, Wafer Lot Acceptance) for further comments.

2.1.2.2.6 MIL-STD-CLSIC (NWSC/ICE): Military Standard for Custom LSI Microcircuits with Design Rules and Quality Requirements.

A. DESCRIPTION

This document represents another effort to generate a military standard for application to custom microcircuits. The document was prepared by Integrated Circuit Engineering (ICE) Corporation under contract to the Naval Weapons Support Center in Crane, Indiana. The final draft is dated June 1982. The stated scope of the document is to "establish the minimum requirements for the manufacture

of standard-design-ruled, LSI-technology, monolithic microcircuits, and the quality and reliability assurance requirements which must be met in the procurement of microcircuits" [punctuation modified]. The proposed MIL-STD includes the use of test structures as process control monitors, referring to NBS Special Publication 400-6 [19] for test chip descriptions. Layout design rules and process control requirements are specified for four specific technologies:

1. Silicon gate NMOS (single and double poly)
2. Silicon gate CMOS (single and double poly)
3. Metal gate NMOS
4. Metal gate CMOS

The appendixes contain tables and flow charts that specify process check points and measurement intervals.

B. COMMENTS

In general, this document relies extensively on MIL-M-38510F, with major differences in only two areas:

1. Wafer acceptance requirements are based on specific electrical and physical parameters utilizing test structures.

2. Design (layout) rules are specified for four MOS technologies.

Specific requirements are given for the use of test devices as process control monitors (PCM). Device parameters for given technologies are prescribed in the appendixes. These parameters are of an electrical nature only, with the exception of Tox, and contain no reliability or geometry predictors. The number of test chips on a wafer is specified as "four or five," and a minimum of three test chips on each wafer must meet the limits defined in the tables. The number of test structures contained in the test chips must be "sufficient" to perform the required process control measurements.

With respect to the custom microcircuit scenario, the following conclusions can be drawn:

1. Since this document relies heavily upon MIL-M-38510F, it also suffers all the same "problems" (e.g., line certification, process inflexibility, and technical deficiencies in some methods and some areas).

2. Requiring device parameter stability through the PCM requirements will aid in the design and fabrication of (relatively) high-yielding chips and help assure wafer fabrication process stability. The PCM requirements do not, however, explicitly address reliability-related factors. To a definite (but unknown) degree, the factors that result in improvement of yield through device parameter stabilization will also improve reliability. However, the capabilities of the specified test structures do not reflect a good comprehension of many major failure

mechanisms. The test chip requirement is thus inadequate, since it falls short of the known capabilities of test structures to identify the presence of major failure mechanisms.

3. The design rule requirements seem to reduce design (and procurement) flexibility. They do have the advantage of providing a conservative design rule baseline. Benefits of this baseline, however, are probably overweighed by the awkwardness of having to fit the specified rules into the framework of a rapidly evolving technology. Unfortunately, the design rules specified in the document are 5-μm rules, which have already been outdated by current fabrication technology. What is needed, in fact, is a set of general requirements (e.g., minimum metal contact overlap) that can be scaled and modified as technological capabilities change.

2.1.2.2.7 MIL-STD-XXXX (RADC/ITT): Quality Procedures for VLSI/VHSIC Type Devices.

A. DESCRIPTION

This documentation represents a currently ongoing effort on the part of Rome Air Development Center (RADC) to update existing military standards and specifications, in particular, MIL-M-38510F and MIL-STD-883C, Method 5007.5 (Wafer Lot Acceptance). The work is being performed under contract to RADC by ITT Advanced Technology Center, LSI Technology Division. The comments given in this report are based solely on the content of ITT Interim Reports Nos. 4 and 5, circa 1984 [11].

B. COMMENTS

The recommendations to RADC by ITT in Report No. 4 generally mirror the current requirements found in MIL-M-38510F, except in the use of test structures for process control and reliability evaluation. The use of test structures also is incorporated into the procedure for line certification. The recommended line certification procedure, under the aegis of an amended 38510, would rely on the utilization of test structures to complement standard evaluations of quality and capability. The goals of these evaluations are to:

1. Verify that the manufacturer's QA program conforms with Appendix A of the amended MIL-M-38510.
2. Verify, in general, that the manufacturer can produce the specified product.
3. Verify, in general, that the manufacturer is capable of sustaining the production of a quality product.

A significant portion of the report contains standardized checklists for use during the evaluation of a manufacturer.

The recommendations contained in Report No. 5 are concerned with modifying wafer lot acceptance requirements (Method 5007, MIL-STD-883C). They constitute a variation of the current Method 5007.5, deleting wafer thickness requirements, tightening thermal stability requirements (to delta VT = 5% of VDD), and adding requirements for line width/pitch and contact resistance stability. The SEM inspection requirements of Method 2018 remain (unfortunately) unchanged. The sampling plan generally relies on one test structure or wafer per lot.

The proposed specification changes appear to suffer from most of the shortcomings of the current MIL-M-38510F, including formal line certification. An effort has been made, however, to utilize new evaluation technology (e.g., test structures) in both line certification and wafer lot acceptance. The proposed wafer lot acceptance requirements are an improvement on 38510F, but are in fact still weak and insufficient in scope. For example, several VLSI failure mechanisms (e.g., TDDB and charge injection,) are not covered at all in the modified 5007. The thermal stability limits, although tighter than in Method 5007.5, are much looser than current technology will support. The work embodied in this document bears following, but the basic approach is fundamentally traditional in nature, and is not likely to result in an optimized qualification/ acceptance specification.

2.1.3 Conclusions

The current military qualification process (represented by the documents reviewed in this report) has proven effective when applied to devices of moderate scale produced by traditional means. It appears, however, that many of the concepts incorporated into this process are ineffective, and many technical requirements either inadequate or inapplicable, when evaluated from the standpoint of custom, foundry-fabricated, VLSI microcircuits.

For example, the Foreword to MIL-STD-976A states that "definite criteria will assure that microcircuits are manufactured under conditions which have been demonstrated to be capable of continually producing highly reliable products." The problems with this premise are at least threefold:

1. The criteria for capability assessment are currently unknown, and are the subject of ongoing research.
2. No such conditions have been demonstrated for devices having the feature sizes typical of "VLSI."
3. Capability does not imply accomplishment. A great deal of attention is currently paid to the examination of the process, the associated quality control apparatus, and the quality of the raw materials. Emphasis should be shifted to an examination of the results of the process by improved testing of the end product.

The current Class S qualification requirements also do not address several major VLSI-level failure mechanisms, and should be modified and strengthened in this regard. Among these failure mechanisms are the following:

1. Time-dependent dielectric breakdown (TDDB).
2. Electrostatic discharge (ESD).
3. Hot carrier effects.

The effects of items 1 and 2, which are specific to CMOS devices, are exacerbated by high density and small feature size. Item 3 is specific to small geometries (less than 3-micrometer feature size).

Furthermore, the current procedures do not reflect the separation of developmental functions (design, wafer fabrication, packaging, etc.) that exist in a "silicon foundry" scenario. In particular, the qualification process must be extended to include the design evaluation through advanced, but proven, analytical techniques. Coverage should include validation and verification of the design at all levels of description, fault simulation, and analyses of performance and testability. MIL-HDBK-339 provides a start in this direction.

Finally, the screening procedures (stressing, testing) to be applied to the final packaged devices need to be strengthened and updated to reflect the wealth of experimental and empirical data that has been gathered over the last several years, and to address the failure mechanisms described above.

It must be emphasized that testing or stressing a component does not make it more reliable. On the contrary, reliability must be an inherent quality of the component: it is a function of the process used to fabricate the device. As a result, the more mature fabrication technologies tend to produce the more reliable devices. The yields of a mature process are also inclined to be higher, although high yields do not necessarily infer high reliability. The extent to which devices are screened, and the severity of the screens, depends both on the type of device and on the intended application [14]. Semiconductor technologies are evolving very rapidly, and maturity of process is becoming increasingly rare. Procedures for qualifying devices produced by those processes must be flexible enough to accommodate this evolution, while providing the user of those devices with well-founded and unambiguous means of ensuring quality and reliability.

The wafer acceptance requirements described in the following section have been excerpted from a set of comprehensive qualification requirements that are currently being developed at JPL in an attempt to satisfy the objectives stated above. This effort, to date, has been conducted under the auspices of the JPL Director's Discretionary Fund as Task Number 288, Qualification Research for Reliable Custom LSI/VLSI Electronics, with additional support from the Product Assurance Technology program regarding wafer acceptance.

2.1.4 Prototype Wafer Acceptance Requirements

2.1.4.1 Preface. The content of this document is based on the assumption that the various qualification issues leading to the formal approval of a custom device for space system usage can be separated into four categories:

1. Supplier qualification
2. Design qualification
3. Process qualification
4. Package qualification

Device qualification, then, is the sum total of these four categories combined with inspection, stressing, and testing of the subject device. Device certification is considered to be the end result of the device qualification process, and applies to individual devices, not to classes or production lots of devices.

The first category, supplier qualification, applies independently to each individual supplier of critical goods and/or services throughout the design, wafer fabrication, packaging, and testing phases of device development and production. The end result of the supplier qualification process is certification of that supplier for a specific type of product or service.

Process qualification refers herein to the evaluation of the wafer processing performed for a specific device, not to the qualification of a particular process for future use. The end result of this process qualification is acceptance of those wafers. The remainder of this prototype requirements document is limited to this qualification stage. The associated procedure is based largely on the extraction of performance and parametric information from specialized test structures, not from measurements made on the subject circuit. These test structure-based measurements can be accomplished quickly and straightforwardly, and can produce results that are both highly accurate and unambiguous. The test structures themselves can be carefully characterized and highly standardized, thus inserting a degree of "user history" into the qualification process that is otherwise missing.

A basic premise of the process qualification requirements is that each wafer must meet specified criteria on an individual, rather than a production lot, basis. The basis for this premise is that few, if any, of the separate fabrication steps to which a wafer is subjected are truly batch processes in the required sense. The processing steps are highly individual in nature, as far as the wafers are concerned, with significant variability of critical parameters between wafers. This individual character is increasingly evident as fabrication techniques become increasingly advanced. Additionally, critical parameters show significant variability across the surface of an individual wafer, a fact which is also reflected in the procedures required by this document.

A further premise is that a circuit designer in a separate organization must be provided with some baseline information on which to predict circuit performance. The requirement resulting from this premise is that the foundry must produce a product that exhibits SPICE parameters within specified tolerances about specified nominal values. The nominal values can relate to those that the foundry can supply readily, but they need to be known to the designer beforehand. The tolerances must, of necessity, relate to the performance tolerances required by the designer in response to the systems application, but may be specified in a generic sense, perhaps by levels related to producibility (and, therefore, yield).

The current content of this document is based on the best information available at the time of writing, and it is intended that the content be updated to reflect the latest state of development, as further investigation separates facts from assumptions. Procedures will be contained in a separate document, and will be based on measurement techniques detailed in JPL Publication 83-70 (Product Assurance Technology for Procuring Reliable, Custom LSI/VLSI Electronics Report for Period: May 1981 - October 1982). Experimental procedures are currently being implemented on custom fabrication runs related to spacecraft systems and programs including CRRES (see Section 2.5) and Mariner Mark II (Digital Filter Processor for X-Band Transponder).

2.1.4.2 Introduction. Criteria for process qualification and wafer acceptance shall be based on the assertion that the role of the foundry is to provide wafers containing a specified minimum number of chips (die) that exhibit a high likelihood of correct functionality, performance, and reliability. The degree of likelihood shall be based on the evaluation of the functionality, performance, and parametric information exhibited or obtained by devices other than the subject circuits, which are located in areas on the wafer closely adjacent to the subject circuits. The criteria, therefore, fall into three categories:

1. Functionality Prediction
2. Performance Prediction
3. Reliability Prediction

This process evaluation shall be performed by means of electrical measurements made on specialized test structures included on the silicon wafers containing the subject circuits and, to some extent, chemical and physical analyses of the materials system of which those wafers are composed. The initial measurements are designed to be quick and qualitative, and to identify regions of the wafers that merit more detailed, quantitative analysis.

2.1.4.3 Criteria

2.1.4.3.1 Functionality Prediction. Two criteria shall deal with predicting the basic functionality of the subject circuit. The first shall be based on the ability of a simple inverter circuit to toggle. An inverter has been chosen as a "canary" circuit since it represents the simplest example of a functional circuit containing both p-channel and n-channel transistors (CMOS technology assumed) and all contact types. The inverter shall be fabricated using the same design rules and feature sizes as the subject circuit, and should provide the quickest indication of potentially correct functionality. More than one inverter shall be provided for interrogation in case one is rendered inoperative by a random defect.

The second criterion shall be based on a measure of the random defect density on the particular wafer. The random defect density (RDD) must be below a specified value for the wafer to be acceptable. Several types of test structures for RDD measurement are currently being evaluated for routine use. Being necessarily large, they would be located in a drop-in test chip (DITC), rather than in an on-chip test strip.

2.1.4.3.2 Performance Prediction. A quick first measure of potential performance shall be provided by the inverter circuit (threshold voltage) and by an adjacent ring oscillator (oscillation frequency/stage delay). These provide the most basic DC and AC parameters, respectively, and are easy and quick to determine. The primary performance prediction criteria shall be based on the demonstrated fabrication of a semiconductor system that yields SPICE parameters within the specified tolerances of the specified nominal values. These parameters shall be provided by extraction from electrical measurements made on a system of specialized test structures located on the DITC. Until such time as these parameters are put on a sound physical basis, this extraction shall be performed by an empirically based computer program such as SUXES (or accepted equivalent).

2.1.4.3.3 Reliability Prediction. Reliability prediction criteria shall be based on a combination of information from specialized test structures located on the DITC and from physical analysis of the materials system contained in the wafer. Electrical and environmental stressing of specified test structures shall also be employed during the final phases of device qualification, after packaging. Verification of adherence to the specified design rules shall be an additional criterion for acceptance from a reliability standpoint, as well as from a performance standpoint. Gross errors, of course, will also affect the basic functionality of the circuit implementations.

2.1.4.4 Requirements. In the following set of requirements, an attempt has been made to minimize the number of test structures, the number of individual tests, and the amount of test equipment required to make those tests. As time progresses, as better correlations are made between the results obtained from test structures and the subject circuits, and as the associated measurement techniques become increasingly refined, a substantial further reduction in time and effort should be realized. Until then, some penalty in time, effort, and silicon area must be tolerated to obtain the necessary information.

2.1.4.4.1 Test Strip. A strip of test structures shall be provided on, or adjacent to, each die. Each test strip shall contain at least two simple inverter circuits having individual access, and a ring oscillator consisting of several (at least five) of these inverters connected in series. Probe pads shall be configured in the standard NBS (2 x N) format. The test strip may be located in the kerf area of the wafer unless it is specified in the procurement document that the integrity of the test strip be preserved in the packaged part.

2.1.4.4.2 Test Chip. A drop-in test chip (DITC) shall be provided in specified sites on each wafer. Each test chip shall contain specialized structures for purposes of device parameter extraction, design rule checking, random fault analysis and reliability analysis. Structures for device parameter extraction shall be adequate in number and type to support SPICE 2G level 3 circuit simulation.

2.1.4.4.3 Test Equipment.

2.1.4.4.3.1 Wafer Handling Equipment. The wafer handling/probing equipment shall be compatible with the standard NBS-developed 2 x N probe-pad array using an 80-μm pad size on a 160-μm grid. The probe card wiring shall support guarded, 4-terminal measurements of required parameters. The probing equipment shall be enclosed to shield the wafer under test from ambient light and to provide a controlled atmospheric environment. Wafer stepping shall be automated to provide repeatable positioning and to facilitate positive identification of the probe location for entry into the database.

2.1.4.4.3.2 Instrumentation. The instrumentation shall support guarded, 4-terminal measurements of the voltage and current, where the forcing voltage or current can be independently controlled. The instrument outputs to the database shall be restricted to representing voltage, current, resistance, conductance, or capacitance. The instrumentation shall be automated to facilitate data acquisition.

2.1.4.4.4 Test Software. This requirement has yet to be established.

2.1.4.5 Wafer Probe Sequence. This sequence of measurements has been designed to allow a quick, qualitative assessment of the viability and uniformity of the wafer, followed by a series of quantitative measurements that are restricted to those regions of the wafer that appear likely to contain acceptable chips.

2.1.4.5.1 Test Strip. The first measurements to be made on each wafer shall be of characteristics of simple circuits ("canaries") contained in the test strips located on each individual chip. The initial test shall be to demonstrate the ability of a simple inverter circuit to toggle. If either inverter toggles correctly, then the threshold voltage is measured and the ring oscillator is interrogated. If neither inverter toggles properly, then the chip is noted as being probably unacceptable and the wafer prober shall proceed to the next chip in sequence without further measurements being made on the "bad" chip. If the threshold voltage of the inverter and the oscillation frequency and power dissipation of the ring oscillator all meet their respective nominal design values within the specified tolerances, then the chip shall be considered to be potentially acceptable. Failing to satisfy any of these criteria establishes the chip as being probably unacceptable.

The "potentially acceptable" and "probably unacceptable" data shall be used to generate a wafer map on which potentially acceptable and unacceptable regions are established. Chips may be selected only from those regions of the wafer where proper inverter and ring oscillator operation are demonstrated, provided that (1) this potentially acceptable region covers at least a specified percentage of the total surface area of the wafer, (2) the observed variation in behavior follows a well-behaved pattern across the wafer surface, and (3) at least one DITC is located within the region. If these three provisos are not met, then the entire wafer is defined as being unacceptable and

further testing is unwarranted except for failure analysis purposes. If all seven of the above criteria are satisfied, then the DITC site(s) is interrogated for more detailed information.

2.1.4.5.2 Test Chip. Following the identification of the potentially acceptable regions of the wafer, parametric measurements shall be made on specialized test structures contained in the DITC(s) located within those regions. Parameters extracted for SPICE simulation shall be within the specified tolerances of their specified nominal values for that region to be considered acceptable. If such parameters fall outside the specified tolerances, the associated region is established as being unacceptable.

Following measurement of the device characteristics for SPICE parameter extraction, other test structures on the DITC shall be interrogated to verify the fidelity of the implementation to the specified design rules. If the nominal values are not met within the specified tolerances, the associated region is established as being unacceptable.

Reliability-related information shall then be obtained from test structures also located on the DITC. When the associated reliability criteria have been satisfied (which may include physical analysis of wafer material), the wafer region shall then be considered to be acceptable. Failing to satisfy these criteria shall establish the region as being unacceptable.

2.1.4.6 Data Acquisition and Analysis. This requirement has yet to be established.

2.1.4.7 Wafer Acceptance Flow Diagram. The wafer acceptance flow diagram is shown in Figure 2.1-1.

2.1.5 References

1. M. G. Buehler, T. W. Griswold, C. A. Pina, B. R. Blaes, C. C. Timoc, R. H. Nixon, and S. F. Suszko, "Product Assurance Technology for Procuring Reliable, Custom LSI/VLSI Electronics," JPL Publication 83-70, Jet Propulsion Laboratory, Pasadena, CA (September 1983).

2. V. Tyree, ISI Test-Chip Site Report, given at the Test Chip Workshop, held at the Jet Propulsion Laboratory, Pasadena, CA, on March 24 and 25, 1983, and subsequent private communication.

3. W. H. Schroen, J. G. Aiken, and G. A. Brown, "Reliability Improvement by Process Control," Reliability Physics Symposium, Las Vegas, NV, 42 (April 1972).

4. J. L. Prince, and B. A. Matsumori, "Qualification Research for Reliable, Custom LSI/VLSI Electronics," Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ (in press).

5. Military Standard: Test Methods and Procedures for Microelectronics, MIL-STD-883C, Department of Defense, Washington, D.C. (August 25, 1983).

6. Military Standard: Certification Requirements for JAN Microcircuits, MIL-STD-976A, Department of Defense, Washington, D.C. (October 5, 1981).

7. Military Standard: Test Methods and Procedures for Microcircuit Line Certification, MIL-STD-977, Department of Defense, Washington, D.C. (January 25, 1982).

8. Military Specification: General Specification for Microcircuits, MIL-M-38510F, Department of Defense, Washington, D.C. (October 31, 1983).

9. Military Handbook: Custom Large Scale Integrated Circuit Development and Acquisition for Space Vehicles, MIL-HDBK-339 (USAF), Department of Defense, Washington, D.C. (July 31, 1984).

10. Integrated Circuit Engineering Corporation, "Military Standard for Custom LSI Microcircuits with Design Rules and Quality Requirements," prepared for Naval Weapons Support Center, Crane, IN (June 4, 1982).

11. ITT Advanced Technology Center, LSI Technology Division, "R & D Status Report Nos. 4 and 5, 'Quality Procedures for VLSI/VHSIC Type Devices'," (including proposed MIL-STD-XXXX, "Certification Requirements for VLSI Microcircuits Facilities and Lines"), prepared for Rome Air Development Center, Rome, NY (1984).

12. A. Dermarderosian, "Water Vapor Penetration Rate Into Enclosures With Known Air Leak Rates," IEEE Trans. Electron Devices, ED-26 (1), 83-90 (January 1979).

13. J. L. Boyle, R. C. McIntyre, R. E. Youtz, and J. T. Nelson, "Latent B-Radiation Damage in Hermetically Sealed NMOS Devices," Reliability Physics Symposium, Orland, FL (1981).

14. E. R. Hnatek, "Effectivity of Burn-in for Integrated Circuits," Institute of Environmental Sciences, Mount Prospect, IL (1980).

15. The Reliability Handbook, National Semiconductor Corporation, Santa Clara, CA, 12 (1982).

16. G. G. Harman, "How the 'New' Technology has Changed and Increased Classical Package-Related Failure Modes," VLSI Packaging Workshop, Las Vegas, NV (1982).

17. L. W. Linholm, "The Design, Testing, and Analysis of a Comprehensive Test Pattern for Measuring CMOS/SOS Process Performance and Control," National Bureau of Standards Special Publication 400-66, pp. 148, U. S. Department of Commerce, National Bureau of Standards, Washington, D. C. (August 1981).

18. "Silicon Foundries: Redefining the Business," VLSI Design, V (8), 24 (August 1981).

19. M. G. Buehler, "Microelectronic Test Patterns: An Overview," National Bureau of Standards Special Publication 400-6, U. S. Department of Commerce, National Bureau of Standards, Washington, D. C. (August 1974).

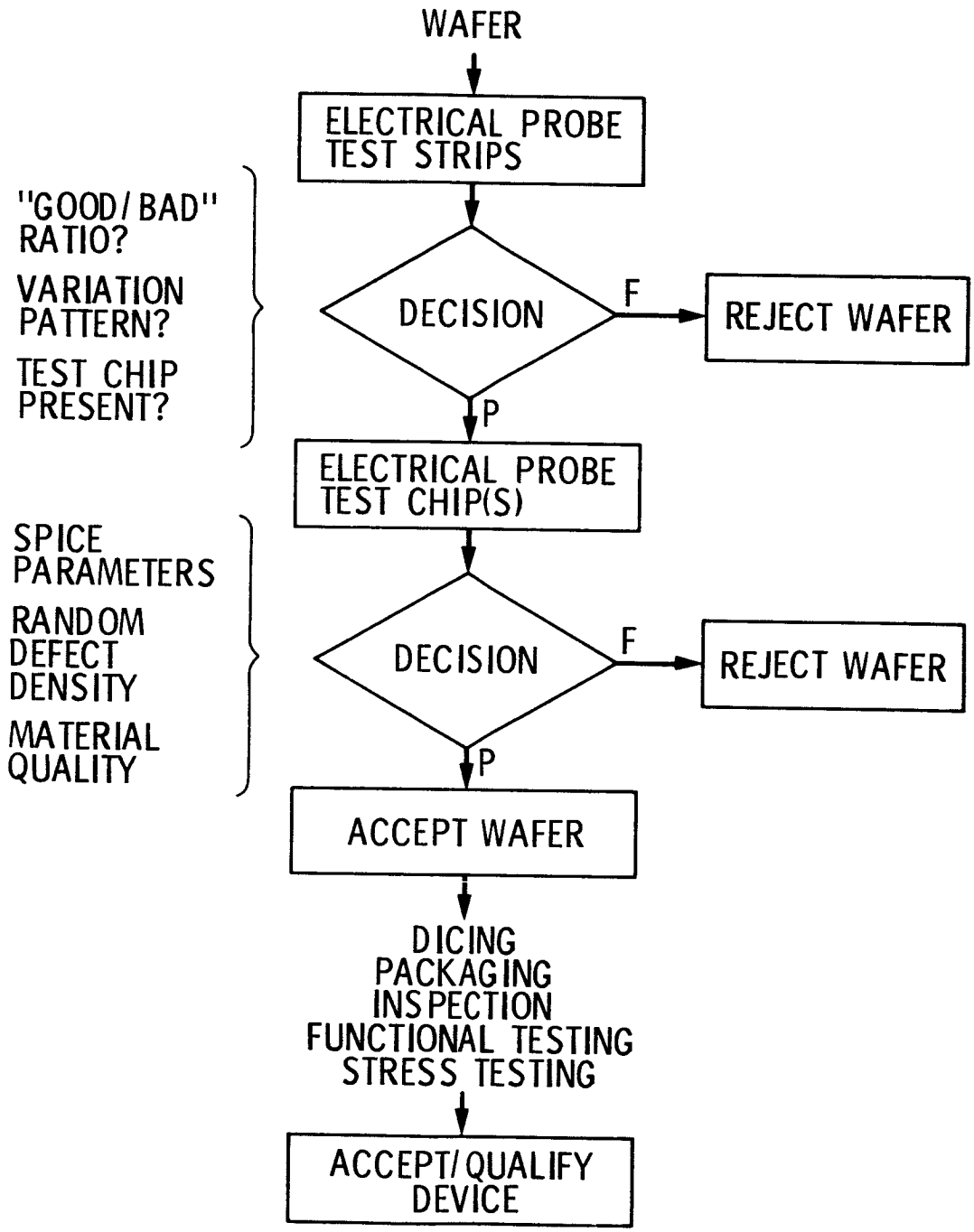


Figure 2.1-1. Wafer acceptance flow diagram

2.2 TEST CHIPS

An integral part of the wafer acceptance procedures described in Section 2.1 is comprehensive test chips. In this section we survey the six categories of the test chips and define the test structures that are included in each category. Then we evaluate the maturity of each test structure and evaluate the area and test time required for each test structure based on our current standard 2-by-N probe pad array and our parametric test system.

2.2.1 CMOS-Bulk Critical Parameter Set[1]

This section lists the parameters that have to be determined to perform wafer or lot evaluations, as well those required to model the behavior of devices and/or circuits in order to predict their performance. This set of parameters was derived for a CMOS process, and the parameters are listed in Table 2.2.1-1, where the body layer refers to either the well or substrate layers. There is some duplication in the listing of the parameters. For example, the layer line-width, the layer sheet resistance, and the contact resistance appear in a number of the categories. The symbols for each test structure are explained in Section 2.2.2. The parameters are listed under the six categories that are described in the following subsections, that were first described elsewhere [1], and that have proven to be a useful classification system.

2.2.1.1 Process Parameters. These parameters are used to monitor the stability of a manufacturing process by measuring those parameters that determine some of the significant process variables such as dopant concentrations, oxide thickness, line-width control of the different layers, and the interlayer contact resistances.

2.2.1.2 Device Parameters. In a CMOS process, the elements of interest are MOSFETs, contacts, and wires. Parameters that characterize these elements are required as inputs to circuit simulation programs. In Table 2.2.1-1, the major parameters required for the SPICE [2] MOSFETs are listed.

2.2.1.3 Circuit Parameters. These parameters are required by the circuit designer to allow the evaluation of noise margins and circuit timing. The majority of these parameters are determined using an inverter or a simple combination of inverters such as ring oscillators.

2.2.1.4 Layout Rules. The information provided by these measurements is required by both the circuit designer and the circuit user. The parameters in this group determine whether or not a circuit can be designed using a given set of geometrical design rules. Although their prime purpose is not that of process control, these parameters can provide important information on the ability of a given manufacturing process to consistently produce devices within a given set of geometrical design rules. The five subcategories listed below follow from the analysis of the layout rules shown in Figure 2.2.1-1. Here it is seen that the layout rules can be described by five uniquely different diagrams: layer width, same layer spacing, different layer spacing, external extensions, and internal extensions. As indicated in the figure, a number of electrical test structures are available to evaluate almost all the rules. In some cases the measurements are direct and provide quantitative numbers, but in other cases the measurements must be indirect and provide qualitative information.

[1]This section was prepared by C. A. Pina and M. G. Buehler.

2.2.1.5 Defect Parameters. The structures in this category are divided between those that can be used for quantitative diagnostic purposes and those that can provide a qualitative indication of the nature of defects. The diagnostic defect density structures provide an indication of the defect density that can be obtained from a wafer lot or individual wafer. The structures used provide a measurement of gate oxide defect density, normally expressed in terms of elements/defect (see Section 2.3.1). Other parameters, such as bridging fault densities or step-coverage faults, can also be determined. The indicator defect density structures are miniature circuits that have row and column addressing. This allows us to pinpoint the location of defects in the circuitry so that visual inspection may be used to ascertain the nature of the defect.

2.2.1.6 Reliability Parameters. The structures associated with this category are currently under development. The potential savings possible after these methods are implemented are judged to be quite large. Two advantages are seen: (a) wafers that exhibit poor reliability parameters can be identified before circuits are packaged, thus saving package and burn-in costs, and (b) overstressing can be performed on test structures that cannot be performed on circuits, thus identifying reliability problems before they show up in the field. Despite the lack of maturity of these structures, the information that can already be obtained with test structures is impressive: (a) the radiation hardness of gate oxides can now be determined at the wafer level by means of a simple electrical test, without the need to actually expose the devices to radiation (see Section 2.3.4.1), and (b) time-dependent oxide breakdown (TDDB) can be used to obtain information on long-term life of gate oxides (see Section 2.3.4.3).

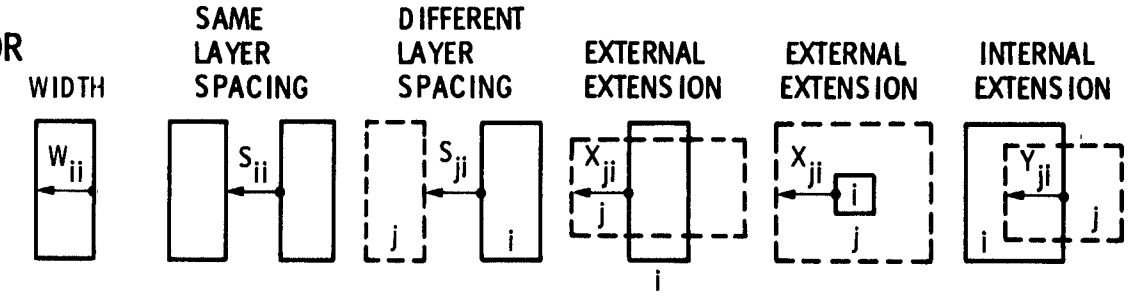
2.2.1.7 References.

1. M. G. Buehler, T. W. Griswold, C. A. Pina, and C. Timoc, "Test Chips for Custom ICs: Six Kinds of Test Structures," Circuits Mfg., 22, 36-42 (June 1982).

2. A. Vladimirescu and S. Liu, "The Simulation of MOS Integrated Circuits Using SPICE2," Electronics Research Laboratory, University of California, Berkeley, Memorandum No. UCB/ERL M80/7 (October 1980).

DIRECT MEASUREMENT

1. SPLIT-CROSS-BRIDGE RESISTOR
2. CAPACITOR



INDIRECT MEASUREMENT

11. ALIGNMENT RESISTOR
12. CONTACT RESISTOR
13. CONTACT RESISTOR ARRAY
14. SERPENTINE RESISTOR
15. COMB RESISTOR
16. COLLISION RESISTOR

W_{ww} = 4 (1)	S_{ww} = 2 (1)	S_{tw} = 5	X_{gt} = 2 (1,2)	X_{tc} = 1 (11)	Y_{tcp} = 3 (12)
W_{tt} = 2 (1)	S_{tt} = 3 (1)	S_{gt} = 1	X_{tg} = 2 (1)	X_{gc} = 1 (11)	
W_{gg} = 2 (1)	S_{gg} = 2 (1,15)	S_{pt} = 2 (16)	X_{pg} = 2 (1)	X_{mc} = 1 (12)	
W_{pp} = 2 (1)	S_{pp} = 2 (1)	S_{gct} = 2	X_{tcp} = 3 (12)	X_{wt} = 3	
W_{cc} = 2 (13)	S_{cc} = 2	S_{pg} = 2 (1)		X_{pt} = 2 (16)	
W_{mm} = 3 (1,14)	S_{mm} = 3 (1,15)	S_{tcg} = 2			

NOTE: RULES IN LAMBDA

T = THIN-OXIDE
W = WELL
G = GATE
P = p + SOURCE/DRAIN DOPING
C = CONTACT
M = METAL

Figure 2.2.1-1. Test structures for checking various CMOS-bulk layout rules

Table 2.2.1-1. Critical Parameters and Associated Test Structures

TEST STRUCTURES PARAMETERS	TEST STRUCTURES[2]
1.0 PROCESS PARAMETERS	
1.1 Layer Line-Width	XBR, SXBR
1.2 Layer Sheet Resistance	XBR, SXBR
1.3 Metal-Layer Contact Resistance	CR, CR-A
1.4 Body Dopant Density	CAP, TR
1.5 Body Resistivity	PFPR
1.6 Layer-Layer Alignment	ALIR
1.7 Body Lifetime	DI, CAP
1.8 Junction Leakage Current (IDBLEAK)	TR
1.9 Junction Breakdown Voltage	TR, DI
1.10 Oxide Thickness	CAP, CAPF
1.11 Gate-Oxide Breakdown Voltage (VBG)	CAP, RTTR
1.12 Field-Oxide Threshold Voltage	TRF
1.13 Channel Leakage Current (IDSO)	TR
1.14 Transistor Punch-Through Voltage (VPT)	TR
2.0 DEVICE PARAMETERS	
2.1 TRANSISTORS	
2.1.1 Threshold Voltage (VTO)	TR, 6TINV
2.1.2 Body Effect Factor (GAMMA)	TR, 6TINV
2.1.3 Conduction Factor (KP)	TR, 6TINV
2.1.4 Effective Channel Width (WE)	TR, 6TINV
2.1.5 Effective Channel Length (LE)	TR, 6TINV
2.1.6 Channel Length Modulation (LAMBDA)	TR, 6TINV
2.1.7 Subthreshold Conduction Factor (NFS)	TR, 6TINV
2.1.8 Gate Oxide Capacitance (COX)	RTTR, CAP
2.1.9 Gate-Source/Drain Overlap Cap. (CG/DSO)	RTTR
2.1.10 Gate-Body Overlap Capacitance (CGBO)	CAPF
2.1.11 Zero-Bias Junction Capacitance (CJ)	RTDI
2.1.12 Zero-Bias Junction Grading Factor (MJ)	RTDI
2.1.13 Zero-Bias Sidewall Junction Cap. (CJSW)	RTDI
2.1.14 Zero-B. Sw. Junc. Cap. Grad. Fact. (MJSW)	RTDI
2.2 CONTACTS	
2.2.1 Contact Resistance	CR
2.3 WIRES	
2.3.1 Layer Sheet Resistance	XBR, SXBR
2.3.2 Layer Line-Width	XBR, SXBR
2.3.3 Layer-Layer Capacitance	STP, CMB
3.0 CIRCUIT PARAMETERS	
3.1 Inverter Output for Input Low (VHIGH)	INV, INV-A
3.2 Inverter Output for Input High (VLOW)	INV, INV-A
3.3 Inverter Threshold for Output = Input (VINV)	NV, INV-A
3.4 Inverter Gain (GAIN)	INV, INV-A
3.5 Inverter Noise Margin (INVNM)	INV, INV-A
3.6 Gate Delay (TAU)	RO, TS

Table 2.2.1-1. (continued)

TEST STRUCTURES PARAMETERS	TEST STRUCTURES[2]
4.0 LAYOUT RULE PARAMETERS	
4.1 Layer Line-Width	XBR, SXBR
4.2 Same Layer Spacing	SXBR, CS
4.3 Different Layer Spacing	CS, ALIR
4.4 Contact Size	CR, CR-A
4.5 Layer Extensions	CS
5.0 DEFECT PARAMETERS	
5.1 DEFECT DENSITY	
5.1.1 Shorts: Different Layers (Pinholes)	CAP-A
5.1.2 Opens: Different Layers (Contact Integrity)	CR, CR-A
5.1.3 Shorts: Same Layer (Layer Bridging)	CMB
5.1.4 Opens: Same Layer (Layer Step Coverage)	STP
5.2 DEFECT LOCATION	
5.2.1 Transistor Matrix	TR-A
5.2.2 Inverter Matrix	INV-A
5.2.3 Static RAM	SRAM
6.0 RELIABILITY PARAMETERS	
6.1 Time-Dependent Dielectric Breakdown	TDDB
6.2 Radiation Hardness	RTR, TR
6.3 Electromigration	CR, CMB
6.4 Oxide Instabilities	TR, CAP
6.5 Contact Integrity	CR, CR-A
6.6 Latch-Up Transistor	LUTR

[2]See Section 2.2.2 for a discussion of each test structure.

2.2.2 Catalog of CMOS Critical Parameters and Test Structures[1]

This section includes a detailed parameter set necessary for wafer and lot evaluation of CMOS integrated circuits using test structures. The parameter set is characterized by area requirements of individual structures, the test and/or evaluation time, the value of the test structure, and an estimate of the structure's maturity status.

In the following material each test structure is described briefly and an abbreviation is assigned to the structure. As seen below, a relatively small number of test structure types are adequate to provide the information for wafer acceptance and device characterization. A majority of the structures have been validated in several technologies and feature sizes: NMOS, CMOS-bulk, and CMOS-SOS, and at feature sizes ranging from 1.2 μm to 5 μm. At the end of this section the structure area considerations are discussed with respect to the use of stand-alone or matrixed test elements.

2.2.2.1 Test Structure Area, Test Time, Value, and Status. The "area" shown per structure depends on whether the test elements in a structure are stand-alone, out-board, or matrixed. For the stand-alone structures the test element is completely surrounded by probe pads. For this case the area is calculated by the rectangle defined by the probe pads. The test elements of each stand-alone structure are connected to an individual probe pad. For these structures we use no common bussing of the structure elements to eliminate the possibility that tests from one structure may affect the results of tests of another.

The out-board structures consist of the round transistors used for hot-carrier studies and ring oscillators. For these structures the area includes the probe pads and the test element. The matrixed structures consist of test element arrays, and the area includes the probe pads and the matrix. For structures contained in a matrix, bussing is, of course, used. For these structures we must consider the possibility that the results are affected by the other elements in the matrix, including the addressing circuitry. The stand-alone or out-board structure approach, rather than the matrixed approach, is felt to be a more satisfactory approach when dealing with wafer/lot acceptance for high-reliability applications.

The following area estimates are based on test structures designed in 3-μm CMOS and the use of the 2-by-10 probe pad array. This array [1] consists of square pads that are 80 μm on a side and spaced 160 μm center-to-center.

The "value" given to each structure is based on a scale of 1 to 5, where a score of 5 indicates the most valued structure. This value system is based on our need at JPL to assess the quality of the wafer fabrication process in terms of parameter means, standard deviations, and defect densities and on our need to acquire parameters for circuit simulation. The structures with a score of 5 are those structures that should be included in any test chip or strip.

[1]This section was prepared by C. A. Pina and B. R. Blaes.

The "test time" given should be considered as a relative time, for it depends on many factors. For instance, the test time will include the data acquisition time, and in some cases it will include some on-line computation and data analysis as well. For example, in the maximum slope method for measuring the MOSFET threshold voltage, an iteration technique is used that requires computer computation to seek out the maximum slope point. In addition, the test time is a function of the particular measurement system. The system in use at JPL consists of a switch matrixed based system that interconnects a set of measuring instruments with a probe card mounted on a wafer prober. The system is controlled with an LSI 11/23 computer. Further details of the system are given elsewhere [2]. All measurements are assumed to be taken in wafer form.

To provide the reader with an indication of the maturity of a test structure, we have provided a status code for each of the test structures. A status of 1 is assigned to a structure if it is in common usage, a status of 2 is assigned if the structure is in use at JPL, and a status of 3 is assigned if the structure is under development at JPL.

2.2.2.2 Critical Test Structures. In the following discussion the suffix A indicates array types of test structures and the suffix F indicates a structure fabricated with a field oxide; otherwise, the structures are fabricated with a gate oxide.

SXBR
XBR

SPLIT-CROSS-BRIDGE RESISTOR [2]. This structure is an eight-terminal resistor that consists of a van der Pauw cross structure to allow the measurement of the layer sheet resistance. In addition, the structure has a bridge resistor and split-bridge resistor to allow the measurement of the line-width and the line spacing. A variation of the split-cross-bridge resistor is the cross-bridge resistor (XBR) [3]. This structure is a six-terminal resistor and so requires an area of only 400 μm x 240 um. The XBR can be used to measure only the sheet resistance and the line-width and not the line spacing. The area, test time, and value for the split-cross-bridge resistor are:

AREA : 560 μm x 240 μm
TEST TIME : 1 s
VALUE : 5
STATUS : 1

CS

COLLISION STRUCTURE [4]. This structure consists of a number of two-terminal structures that are designed to evaluate certain layout rules that cannot be evaluated by another structure, such as the overlap of the doping layer beyond the diffusion layer and the overlap of the metal beyond the contact window. These structures consist of a number of individual segments in which one part of the individual segments is varied geometrically. An open or short condition indicates when a geometrical boundary has been exceeded.

AREA : 80 μm x 240 μm
TEST TIME : 1 s
VALUE : 3
STATUS : 3

PFPR — PLANAR-FOUR PROBE RESISTOR [5]. This structure has four terminals all connected to the body layer. It is used to determine the body resistivity using DC current and voltage measurements.

AREA : 240 μm x 240 μm
TEST TIME : 1 s
VALUE : 2
STATUS : 1

ALIR — ALIGNMENT RESISTOR [6]. This structure has ten terminals and is used to determine the alignment between such layers as metal-to-contact and metal-to-diffusion using DC current and voltage measurements.

AREA : 720 μm x 240 μm
TEST TIME : 2 s
VALUE : 3
STATUS : 1

CR — CONTACT RESISTOR [7]. This structure is a four-terminal resistor designed to measure the contact resistance between two different layers. By using a selected group of contact resistors of varying areas, a measure of the process quality (i.e., the "process cliff") can be obtained (see Section 2.3.2).

AREA : 240 μm x 240 μm
TEST TIME : 0.5 s
VALUE : 5
STATUS : 1

CR-A — CONTACT RESISTOR ARRAY. This structure has 20 contacts that include connections to addressing circuitry and analog inputs to the contacts in the array. The structure contains 225 individually addressable contacts of various sizes to determine contact variability and to determine the contact resistance cliff. Contacts in the array are accessed by means of a binary addressing scheme.

AREA : 1520 μm x 1520 μm for 225 CR
TEST TIME : 150 s (~0.8 s/device)
VALUE : 3
STATUS : 3

DI — DIODE. This structure has three terminals with connections to the source/drain diffusion, body, and peripheral gate. This structure is a large area structure that requires the space of four probe pads so that peripheral effects can be minimized. It is used to measure junction leakage currents, breakdown voltage, and lifetime. The structure is intended to be large enough so that bulk leakage currents can be easily measured.

AREA : 240 μm x 240 μm
TEST TIME : 1 s
VALUE : 4
STATUS : 1

RTDI — CLOSED-GEOMETRY DIODES (Race Track Diodes). This structure has two terminals with connections to the source and body. In contrast with the DI, this structure has no peripheral gate. This structure is designed to have the same shape as the RTTR so as to allow the measurement of the junction peripheral capacitance. Thus this structure requires the space of 10 probe pads.

AREA : 720 μm x 240 μm
TEST TIME : 1 s
VALUE : 4
STATUS : 2

TR
TRF — TRANSISTOR. This structure is a four-terminal MOSFET with connections to the source, drain, gate, and body. The test time for the TRF is much shorter than for the TF, as indicated below.

AREA : 240 μm x 240 μm
TEST TIME : 5 s (TR), 0.5 s (TRF)
VALUE : 5
STATUS : 1

RTR — ROUND TRANSISTOR. This structure has four terminals with connections to the source, drain, gate, and body where the drain completely encloses the source. Thus the gate is contacted over thin oxide. The round transistor is used for hot carrier injection into gate oxides to evaluate the radiation hardness of gate oxides (see Section 2.3.4.1).

AREA : 240 μm x 350 μm
TEST TIME : 15 s
VALUE : 4
STATUS : 3

RTTR — CLOSED-GEOMETRY TRANSISTORS (Race Track TRansistors). This has four terminals with connections to the source, drain, gate, and body. Thus the gate is contacted over thin oxide. The structure requires the space of 10 pads so as to aid in the measurement of parasitic gate overlap capacitance and gate oxide thickness from capacitance measurements.

AREA : 720 μm x 240 μm
TEST TIME : 1 s
VALUE : 4
STATUS : 2

TR-A — TRANSISTOR ARRAY. This structure has 20 terminals with connections to VDD, GND, addressing circuitry, and the analog inputs to the transistors. The transistor array can accomodate 16,384 transistors and is designed using minimum geometry design rules. Thus it is useful for evaluating transistor parameter variations and processing defects.

AREA : 1520 μm X 1520 μm
TEST TIME : ?
VALUE : ?
STATUS : 3

LUTR — LATCH-UP TRANSISTOR [8]. This structure has four terminals with connections to the emitter, base, well, and substrate.

AREA : 240 μm x 240 μm
TEST TIME : 1 s
VALUE : 3
STATUS : 3

INV — INVERTER. This structure has four terminals with connections to the input, output, VDD, and GND. For this structure the testing can be very simple and can consist of only five voltage points: VHIGH, VLOW, VINV, VINV + 25 mV, and VINV - 25 mV. The test time given below is for this five-point set.

AREA : 240 μm x 240 μm
TEST TIME : 1 s
VALUE : 5
STATUS : 1

6TINV — SIX-TERMINAL INVERTER. This structure has six terminals with connections to the input, output, VDD, GND, well, and substrate. The two additional terminals, well and substrate, allow a full characterization of each individual transistor as well as the inverter. This structure is under development at the present time.

AREA : 400 μm x 240 μm
TEST TIME : 10 s
VALUE : ?
STATUS : 3

INV-A — INVERTER ARRAY [9]. This structure has 20 contacts that include connections to addressing circuitry and analog inputs to the inverters in the array. The structure contains 222 individually addressable inverters to determine variability in the inverter transfer curves. Inverters in the array are accessed by means of a shift register addressing scheme (see Section 2.3.3).

AREA : 1520 μm x 1520 μm
TEST TIME : 300 s (~1.4 s/device)
VALUE : 3
STATUS : 3

RO — RING OSCILLATOR. This structure has five contacts that include the VDD(inverter chain), VDD(output amplifier), GND, output, and trigger. The number of inverter stages is a prime number and is usually between 19 and 31 stages. This structure does not fit within the 2-by-10 probe pad array and hence is designed to fit off one end of the pad array.

AREA : 1120 μm x 240 μm for xxx inverters
TEST TIME : 0.5 s
VALUE : 4
STATUS : 1

TIMING SAMPLER. This structure consists of a chain of 128 inverters that are connected to latches which in turn are connected to a decoder circuit. It has ten terminals with connections to VDD, GND, input, enable, and decoded output (see Section 2.6.3).

AREA : 3800 μm x 720 μm
TEST TIME : 0.5 s
VALUE : 4
STATUS : 2

SRAM

STATIC RAM. This structure, as implemented as a 16-by-64 RAM, has 28 terminals with connections to the VDD, GND, address bus, data bus, and control logic (see Section 2.6.1).

AREA : 1520 μm x 1800 μm
TEST TIME : 5 s
VALUE : 4
STATUS : 2

CAP
CAPF

CAPACITOR. This structure is a MOS capacitor that has two terminals with connections to the gate and body. The structure is used to determine the gate-oxide capacitance, body dopant density, and body lifetime using capacitance-voltage measurements. To minimize the effects of peripheral capacitance, the structure is a large area structure that requires the space of 10 probe pads.

AREA : 720 μm x 240 μm
TEST TIME : 5 s
VALUE : 5
STATUS : 1

CAP-A

CAPACITOR ARRAY. This structure consists of a metal cap over a network of transistors formed by poly fingers orthogonally crossing diffusion fingers. The structure has 10 contacts that include connections to the metal, gate subarrays, diffusion, and body. Pairs of pads are connected together to allow for a probe down test. This structure is used to determine the frequency of faults between the metal and poly and between the poly and silicon using DC leakage current measurements (see Section 2.3.1).

AREA : 1520 μm x 1520 μm for 50,xxx transistors
TEST TIME : 5 s
VALUE : 3
STATUS : 2

TDDB

TIME-DEPENDENT DIELECTRIC BREAKDOWN STRUCTURE. This structure has 40 probe pads and consists of a network of transistors formed by poly fingers orthogonally crossing diffusion fingers. This structure requires a large area and long test times in order to characterize the breakdown statistics.

AREA : 3040 μm x 6080 μm
TEST TIME : 600 s = 10 min
VALUE : 2
STATUS : 3

STP CMB — STEP and COMB ARRAY. This structure consists of two interdigitated combs with a serpentine tracing the space between the combs. The structure is fabricated in a conducting layer (usually metal or poly) and is designed with a number of subarrays. The structure has connections to the layer subarrays, the crossing layer if poly, and the body. Pairs of pads are connected together to allow for a probe down test. This structure is used to determine the frequency of shorts between the combs and the serpentine and the frequency of opens in the serpentine. In the latter case, the structure is used as a step-coverage test. The faults are evaluated by measuring the leakage current between the elements of the structure.

AREA : 1520 μm x 1520 μm
TEST TIME : 1 s
VALUE : 3
STATUS : 3

2.2.2.3 Compact Test Structures. The area requirements shown above assumed that the stand-alone test structures have one probe pad dedicated to each terminal in the structure. The matrixed (or compact) test structures, of course, share probe pads. In Table 2.2.2-1 we compare the space savings for the matrixed structures over the stand-alone structures. As seen in the table, a considerable area savings can be effected by using matrixed test structures. But, as mentioned above, the matrixed structures have the disadvantage that the additional circuitry can reduce confidence in the results; however, the matrixed structures are a good approach when statistical information on parameter variations is being acquired and when it is necessary to pinpoint the location of defects.

Assuming a four-inch diameter wafer with approximately 90 prime sites, 10% of the wafer area is required for unbussed test structures to obtain parametric and defect density information. If high-reliability information is desired, then an additional 10% to 20% of the wafer area may be required (for a total of 20% to 30% of the wafer area). The use of compact test structures should significantly reduce the area requirements, or, alternatively, provide additional statistics using the same area.

2.2.2.4 References.

1. M. G. Buehler, "Comprehensive Test Patterns with Modular Test Structures: The 2 by N Probe-Pad Array Approach," Solid State Technol., 22, 89-94 (October 1979).

2. M. G. Buehler, T. W. Griswold, C. A. Pina, B. R. Blaes, C. C. Timoc, R. H. Nixon, and S. F. Suszko, "Product Assurance Technology for Procuring Custom LSI/VLSI Electronics," JPL Publication 83-70, Pasadena, California (September 1983).

3. M. G. Buehler, S. D. Grant, and W. R. Thurber, "Bridge and van der Pauw Sheet Resistors for Characterizing the Line Width of Conducting Layers," J. Electrochem. Soc., 125, 650-654 (1978).

4. B. M. M. Henderson, A. M. Gundlach, and A. J. Walton, "Integrated-Circuit Test Structure Which Uses a Vernier to Electrically Measure Mask Misalignment," Electr. Lett., 19, 868-869 (1983).

5. M. G. Buehler and W. R. Thurber, "A Planar Four-Probe Test Structure for Measuring Bulk Resistivity," IEEE Trans. Electron Devices, ED-23, 968-974 (1976).

6. T. J. Russell and D. A. Maxwell, "A Production-Compatible Microelectronic Test Pattern for Evaluating Photomask Misalignment," NBS Special Publication 400-51 (April 1979).

7. S. J. Proctor and L. W. Linholm, "Direct Measurement of Interfacial Contact Resistance, End Contact Resistance, and Interfacial Contact Layer Uniformity," IEEE Trans. Electron Devices, ED-30, 1535-1542 (1983).

8. A. Ochoa and P. V. Dressendorfer, "A Discussion of the Role of Distributed Effects in Latch-Up," IEEE Trans. Nucl. Sci., NS-28, 4292-4294 (1981).

9. M. G. Buehler and H. R. Sayah, "Addressable Inverter Matrix for Process and Device Characterization," Solid State Technol., 28, 185-191 (May 1985).

Table 2.2.2-1. Comparison of the Area per Element for Stand-Alone (TR and INV) and Matrixed (TR-A and INV-A) Test Structures

TEST STRUCTURE	AREA (μm^2)	ELEMENTS/STRUCTURE	AREA/ELEMENT (μm^2/ELEMENT)	RATIO
TR	240 x 240	1	57,600	408
TR-A	1520 x 1520	16,384	141	
INV	240 x 240	1	57,600	6
INV-A	1520 x 1520	222	10,407	

2.3 TEST STRUCTURES

The test structures discussed in this section were selected to advance the state of the art in defect detection and yield analysis (Pinhole Array Capacitor), in contact parameter spread analysis (contact resistance process cliff), in inverter parameter spread analysis (addressable inverter matrix), in device reliability (reliability structures), in worst-case circuit design (inverter noise margin analysis), and in MOSFET parameter extraction (JMOSFIT).

2.3.1 Pinhole Array Capacitor[1]

During the contract period, the Pinhole Array Capacitor (PAC), described elsewhere [1], was further developed. A detailed testing procedure was developed, results were analyzed from data obtained from 1.2-μm CMOS PACs, and fault models were developed for gate-oxide shorts.

2.3.1.1 Test Procedure. The testing procedure for the PAC consists of three tests. The first test is a probe check that identifies the location of the pad-pair where probes have not made good contact with the pads. The second test evaluates metal-poly shorts. The third test evaluates polysilicon gate-oxide shorts. These three tests enable the extraction of yield and fault modeling information.

A transistor-level description for the PAC is shown in Figure 2.3.1-1. In the following tests the pads that are not included in the test are allowed to float electronically. The time allowed for each current measurement is 0.1 seconds. The test procedures for each of the three tests are given in the following subsections.

2.3.1.1.1 Probe Check Test. Measure the two-terminal resistance between pad pairs 1/20, 2/19, 3/18, . . . , 10/11. If the measured resistance is greater than a resistance threshold value (we used 1 megohm), then set a flag for that pad-pair. Any subsequent measurements obtained through a flagged pad-pair will be marked invalid.

2.3.1.1.2 Metal-Poly Oxide Short Test. Currents are measured between the pad-pairs indicated below and ground:

Subarray 1 - Apply voltage (+5 V) to pads 1/20, then measure the current from pads 2/19.

Subarray 2 - Apply voltage (+5 V) to pads 3/18, then measure the current from pads 4/17.

Subarray 3 - Apply voltage (+5 V) to pads 5/16, then measure the current from pads 6/15.

Subarray 4 - Apply voltage (+5 V) to pads 7/14, then measure the current from pads 8/13.

[1]This section was prepared by B. R. Blaes and H. R. Sayah.

If the measured current is greater than or equal to the cutoff current, I(CUTOFF), then the subarray is considered to have a metal-poly short.

2.3.1.1.3 Gate-Oxide Short Test. The test for the gate-oxide shorts requires four current measurements: ID(ON), IB(ON), ID(OFF), and IB(OFF). ID is the current from the poly to the diffusion, and IB is the current from the poly to the bulk. (ON) means that the channel is present, while (OFF) means that the channel is absent. For n-type arrays, +5 volts are used to turn on the channel, and -5 volts are used to turn off the channel. For p-type arrays, the voltage signs are reversed. When ID is being measured, pads 9/12 are grounded, and when IB is being measured, pads 10/11 are grounded.

Subarray 1 - Apply voltage ON and OFF between pads 2/19. Ground and measure the current ID(ON), and ID(OFF) between pads 10/11. Ground and measure the current IB(ON), and IB(OFF) between 9/12 and ground.

Subarray 2 - Apply voltage ON and OFF between pads 4/17. Ground and measure the current ID(ON), and ID(OFF) between pads 10/11. Ground and measure the current IB(ON), and IB(OFF) between 9/12 and ground.

Subarray 3 - Apply voltage ON and OFF between pads 6/15. Ground and measure the current ID(ON), and ID(OFF) between pads 10/11. Ground and measure the current IB(ON), and IB(OFF) between 9/12 and ground.

Subarray 4 - Apply voltage ON and OFF between pads 8/13. Ground and measure the current ID(ON), and ID(OFF) between pads 10/11. Ground and measure the current IB(ON), and IB(OFF) between 9/12 and ground.

The nature of the pinhole can be determined by analysis of the four currents: ID(ON), IB(ON), ID(OFF), and IB(OFF). (See Section 2.3.1.3.)

2.3.1.2 Test Results from 1.2-μm CMOS PACs. The number of elements designed into each of the subarrays in the 1.2-μm PAC is shown in Figure 2.3.1-1. The total number of elements for the structure is 136,111, and the structure required an area 1.2 mm by 1.7 mm. A yield analysis [1] for 14 n-channel PACs per wafer is shown in Figure 2.3.1-2 and for 28 p-channel PACs per wafer in Figure 2.3.1-3. From the analysis, the characteristic number for the n-channel PAC is E = 61,500 elements/defect and for the p-channel PAC E = 78,000 elements/defect. These values were determined for I(CUTOFF) = 10 nA.

2.3.1.3 Gate-Oxide Fault Models. The PACs were fabricated using a CMOS local oxidation process in which silicon nitride is used to define the gate oxide. The gate-oxide pinholes are believed to have their origin where a residual nitride remains at the silicon surface which masks against gate-oxide growth [2, 3, 4]. This residual nitride layer results in the thinning of the gate oxide at the affected regions (see Figure 2.3.1-4).

As explained briefly in Section 2.3.1.1, the four PAC current measurements are used to determine the nature of the pinhole defect. The defect classes are listed in Table 2.3.1-1. As can be seen in Figure 2.3.1-5, two types of PAC fault models are proposed to explain the four measured currents (ID(ON), IB(ON), ID(OFF), and IB(OFF)). The type #1 defect models the n-channel PAC, and the type #2 defect models the p-channel PAC. In the first type, the pinhole forms an ohmic connection to the channel and a diode connection to the bulk. In the second type, the pinhole forms a diode connection to the channel and an ohmic contact to the bulk. From these models we have prepared an expected response for the currents, as shown in Table 2.3.1-2. If the measured current is greater than or equal to I(CUTOFF), then it is assigned the value "1." Otherwise it is assigned the value "0." From the four measured currents one can identify the nature of the defect. In some cases, other combinations of currents are observed. In these cases, the defect is not modeled by the defects shown in Figure 2.3.1-5. In such cases, the defects usually cover a large area and affect other, adjacent subarrays. We are in the process of collecting the data that will confirm the predictions of Table 2.3.1-2.

2.3.1.4 References.

1. M. G. Buehler, B. R. Blaes, C. A. Pina, and T. W. Griswold, "Pinhole Array Capacitor for Oxide Integrity Analysis," Solid State Technol., 26, 131-137 (November 1983).

   Note: This article is reprinted in its entirety as it appeared in the journal, following Table 2.3.1-2.

2. E. Kooi, J. G. van Lierop, and J. A. Appels, "Formation of Silicon Nitride at a Si-SiO_2 Interface During Local Oxidation of Silicon and During Heat-Treatment of Oxidized Silicon in NH_3 Gas," J. Electrochem. Soc., 123, 1117-1120 (1976).

3. O. Nakajima, N. Shiono, S. Musamoto, and C. Hushimoto, "Defects in a Gate Oxide Grown After the LOCOS Process," Jpn. J. Appl. Phys., 18, 943-951 (1979).

4. C. A. Goodwin and J. W. Brossman, "MOS Gate Oxide Defects Related to Treatment of Nitride Coated Wafers Prior to Local Oxidation," J. Electrochem. Soc., 129, 1066-1070 (1982).

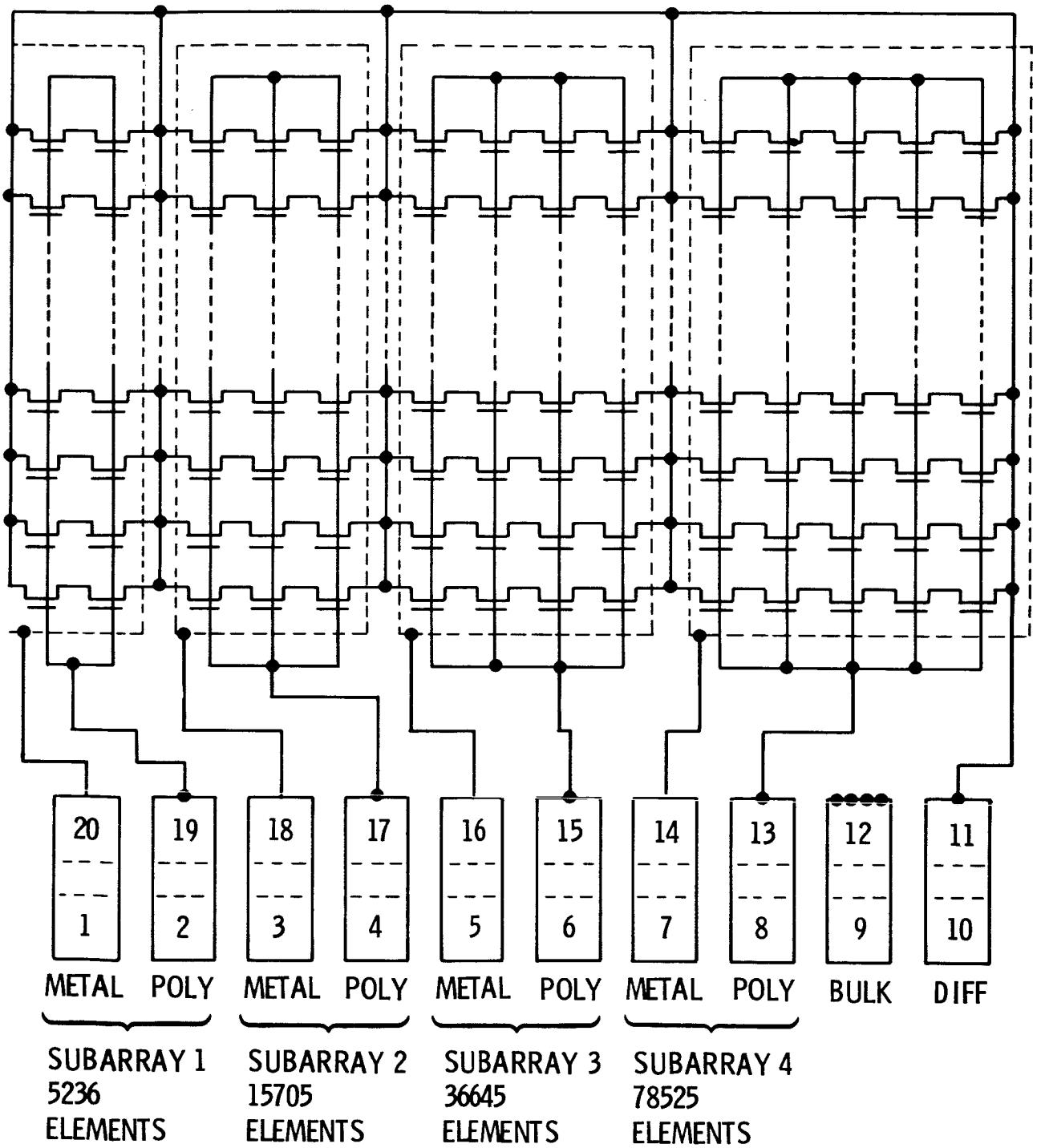


Figure 2.3.1-1. Transistor-level description of the Pinhole Array Capacitor

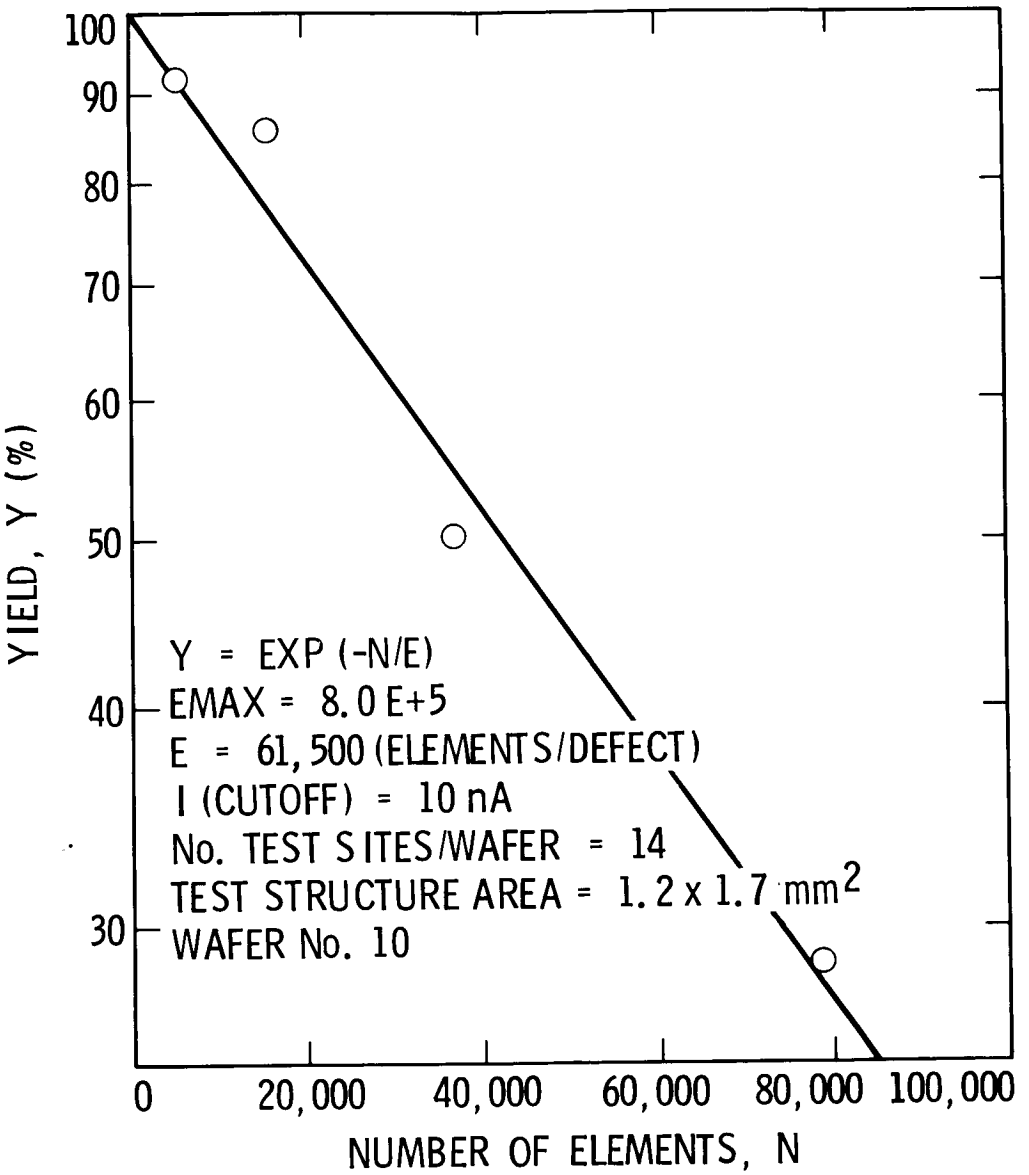


Figure 2.3.1-2. Poly-bulk shorts for an n-channel Pinhole Array Capacitor fabricated with a 1.2-μm CMOS process

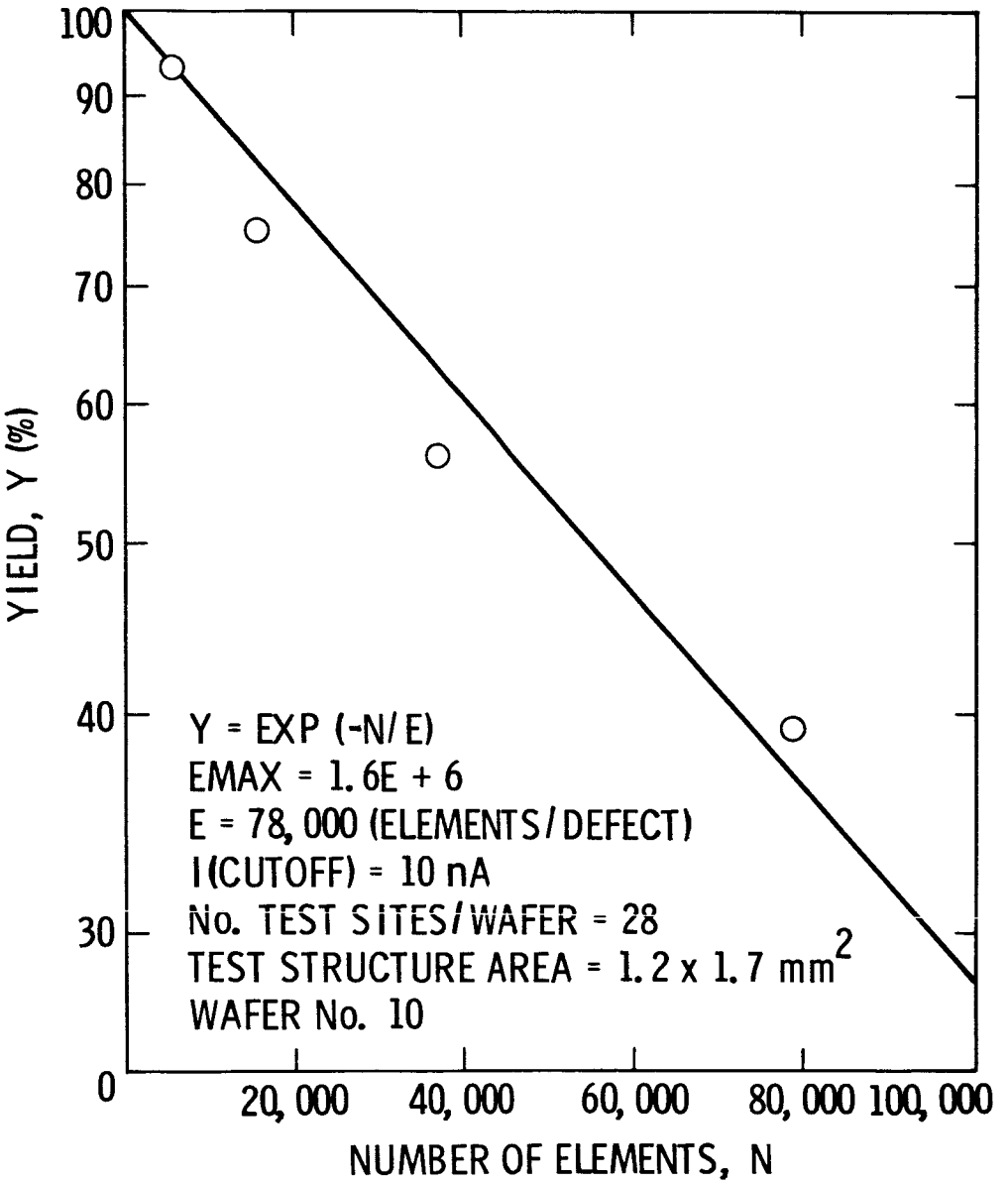


Figure 2.3.1-3. Poly-bulk shorts for a p-channel Pinhole Array Capacitor fabricated with a 1.2-μm CMOS process

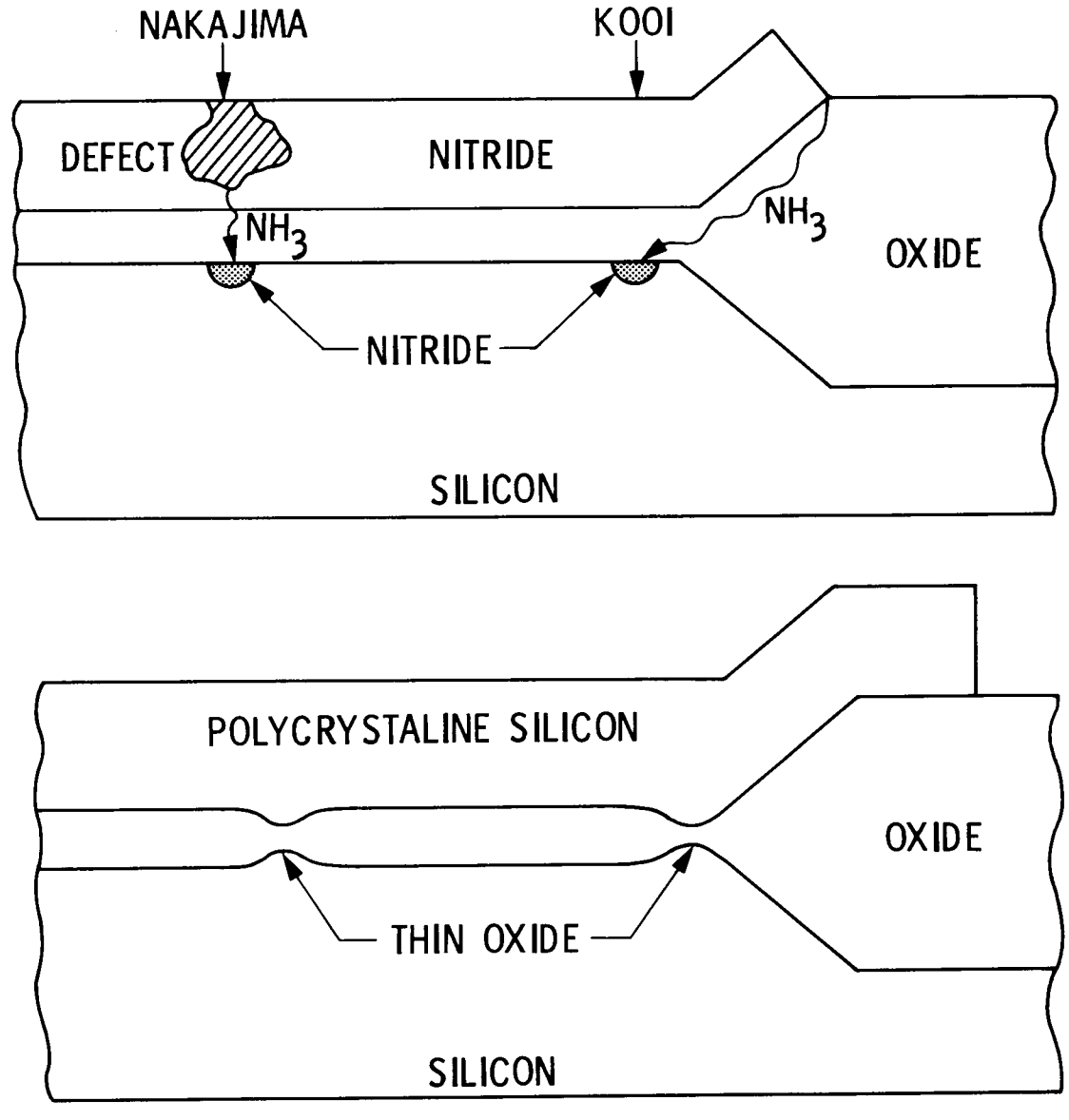


Figure 2.3.1-4. Origin of oxide pinholes in a CMOS local oxidation process where silicon nitride is used to define the thin oxide regions. After Kooi [2] and Nakajima [3].

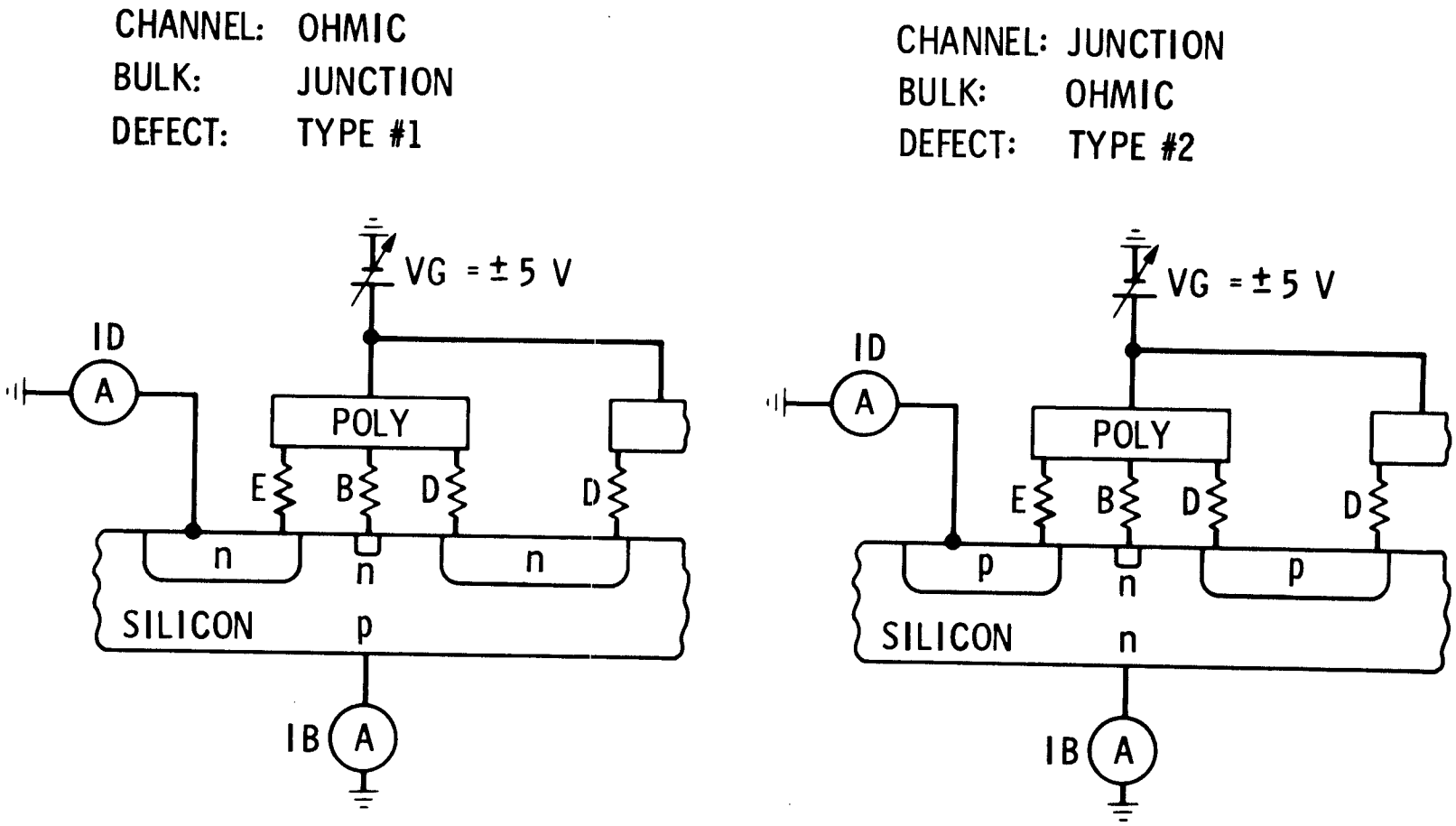


Figure 2.3.1-5. Two types of Pinhole Array Capacitor defects

Table 2.3.1-1. Pinhole Array Capacitor Defect Classes

DEFECT CLASSES
N = NO DEFECT DETECTED
B = POLY-BULK DEFECT
D = POLY-DIFFUSION DEFECT
E = POLY-DIFFUSION DEFECT (EDGE OF ARRAY)
? = DEFECT NOT MODELED

Table 2.3.1-2. Defect Identification Based on Four Pinhole Array Capacitor Tests

ID(ON)	IB(ON)	ID(OFF)	IB(OFF)	TYPE 1	TYPE 2
0	0	0	0	N	N
1	0	0	1	B, D, DB	D
1	0	1	1	E, EB, ED, EBD	E, ED
1	1	0	1	?	B, BD
ALL OTHER COMBINATIONS				?	?, EB, EBD

ON = Channel present
OFF = Channel absent

0 = Current less than I(CUTOFF)
1 = Current greater than I(CUTOFF)

Pinhole Array Capacitor for Oxide Integrity Analysis*

M. G. Buehler B. R. Blaes C. A. Pina T. W. Griswold
Jet Propulstion Laboratory, California Institute of Technology, Pasadena, California

Reprinted from Vol. 26, No. 11
SOLID STATE TECHNOLOGY

The integrity of the metal-poly oxide and the gate oxide was evaluated for several 5-μm CMOS-bulk processes. The pinhole array capacitor consists of diffused and poly fingers that form a network of MOS transistors (elements), which are capped by a deposited oxide and metal layer. The smallest structure used in this study contained about 15,000 elements and the largest structure contained about 68,000 elements. Each structure was divided into several subarrays. The structures are place a number of times on each wafer. From a yield analysis of the subarrays, the elements per defect were found to be typically in excess of 50,000 elements/defect for the metal-poly oxide and 100,000 elements/defect for the gate oxide. From the switching behavior of the transistors, the gate oxide defects were tentatively identified as gate-to-body shorts rather than gate-to-diffusion shorts.

SHORTS BETWEEN CONDUCTING LAYERS have been observed to contribute noticeably to the yield loss in MOS circuits [1-3]. Such faults can lead to a variety of circuit malfunctions including stuck-at faults for low impedence shorts and slow-to-rise or slow-to-fall for high impedence faults [4]. The objective of this effort was to develop a pinhole array capacitor (PAC) that can be fabricated along with integrated circuits, and to characterize the frequency of occurrence of oxide defects such as pinholes and nonlocal defects. Ultimately we hope to assess the importance of this physical failure mechanism relative to other failure mechanisms such as open contacts, open metal at oxide steps, and shorts between like conductors [5].

The challenge in developing the PAC was found in designing a structure that conserves area on a wafer and yet allows one to adequately characterize the oxide pinhole density. The PAC was designed to be very compact by designing features at the layout rule limits. Four different test structures were designed and fabricated in the course of this study, and they varied from the smallest with 14,994 elements to the largest with 68,328 elements. The challenge in analyzing the data was found in sorting randomly occurring pinholes from photomask flaws and nonlocal defects such as scratches, and in identifying the nature of gate-oxide defects.

*This effort was sponsored jointly by the National Aeronautics and Space Administration, the Defense Advanced Research Projects Agency, and the National Security Agency under contract number NAS-918.

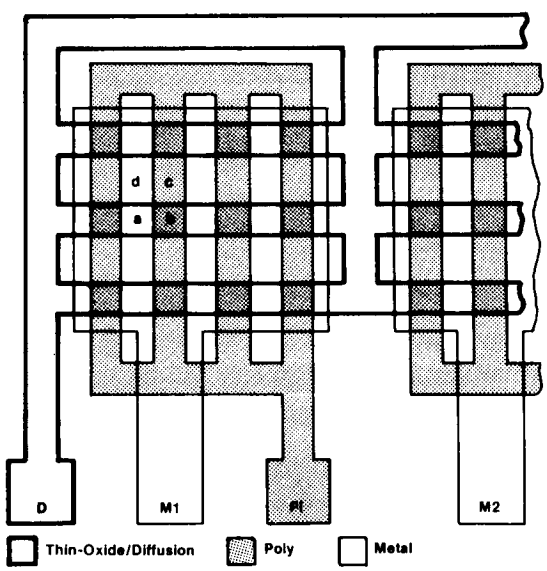


Fig. 1—Schematic layout of the pinhole array capacitor, where an "element" is represented by the region labeled a-d. The structure also contains a contact to the body which is a p-well for an n-channel structure or an n-substrate for a p-channel structure.

Test Structure

Various designs for pinhole capacitors may be found in the literature [2, 6, 7]. The pinhole array capacitor used in this study is shown in Fig. 1 and was fabricated using several

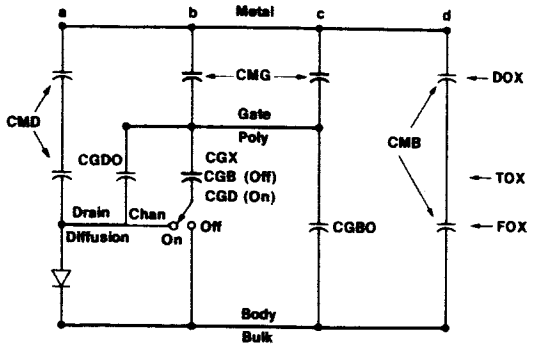


Fig. 2—Equivalent circuit diagram of pinhole array capacitor, where a-d refer to points shown in Fig. 1. ON and OFF refer to the respective presence or absence of the inversion channel under the poly gate.

5-μm CMOS-bulk p-well processes [8]. The fabrication was begun by forming a network of thin-oxide fingers using an isoplanar process. After a thermal gate oxidation growth, a phosphorus-doped polycrystalline silicon (poly) layer was patterned into a network of fingers that cross the thin-oxide fingers orthogonally. In the self-aligned process used to fabricate this structure, the thin-oxide regions between the poly fingers were doped to form transitors. An oxide layer was then deposited and the structure was capped with a metal (aluminum) layer. Finally, a glassy layer was deposited to protect the structure. The structures described in this paper were designed using the 5-μm CMOS-bulk layout rules described elsewhere [9].

The structure consists of a network of MOS transistors whose channels are controlled by the gate-to-body potential. An equivalent circuit is shown in Fig. 2, where the three kinds of insulators are identified: namely, the thin (or gate) oxide (TOX), the field oxide (FOX), and the deposited oxide (DOX). Capacitor CGX is connected to a switch that connects the capacitor to the bulk (gate-body) capacitor (CGB), when the transistor channel is off, and to the channel and hence to the diffusion (gate-diffusion capacitor, CGD) when the transistor channel is on. The bulk is either the p-well for an n-channel structure or the n-substrate for a p-channel structure.

The capacitors most susceptible to pinholes are shown in Fig. 2 with bold lines. They are the metal-gate capacitor, CMG, and the CGX capacitors. This knowledge guided our experimental measurements and reduced the number of possible leakage current measurements to four. The other capacitors shown in Fig. 2 are the metal-diffusion capacitor, CMD; the gate-diffusion overlap capacitor, CGDO; the gate-body overlap capacitor, CGBO; and the metal-body capacitor, CMB.

The overall view of the PAC is seen in the photomicrographs shown in Figs. 3 and 4. As seen in these figures, each structure is divided into a number of subarrays. The number of elements in each subarray is listed in Table I. An element is defined as the region labeled a-d in Fig. 1.

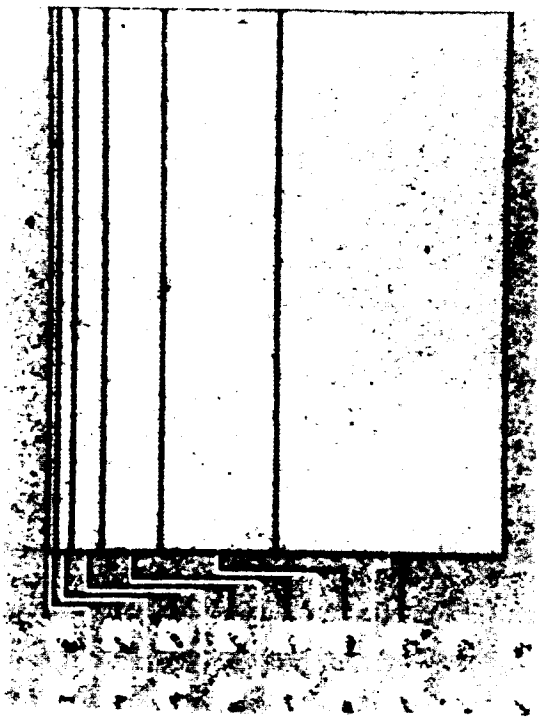


Fig. 3—Photomicrograph of pinhole array capacitor No. 1.

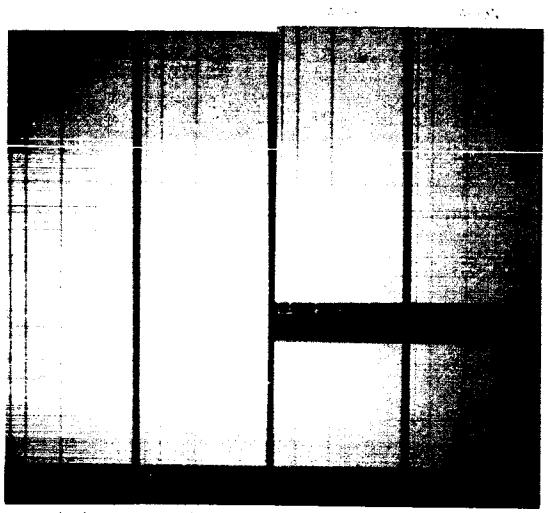


Fig. 4—Photomicrograph of pinhole array capacitors No. 2–4 with double probe pads that facilitate a probe down test.

The PAC was designed to be probed using a 2 by 10 probe array [10]. Two probing arrangements were used, and these are indicated in Table I. The number of elements in a particular subarray is listed between the metal and poly probe

Table I—Pinhole Array Capacitor Details

Run No.	Test Chip	Test sites per Wafer, N_s	E_{MAX} (Elements/Defect)	Array No.	Probe Pad Configuration No. Elements per subarray, N_i									
					D	P1	P2	P3	P4	P5	P6	X	X	X
					M1	M2	M3	M4	M5	M6	X	X	X	X
1	2012	68	351,000	PAC-1		238	476	952	1904	3808	7616			
7	3012	32	165,000	PAC-1		238	476	952	1904	3808	7616			
					M1	P1	M2	P2	M3	P3	M4	P4	SW	D
					M1	P1	M2	P2	M3	P3	M4	P4	SW	D
8	3054	13	107,000	PAC-2	756		2268		6292		11340			
8	3054	13	237,000	PAC-3	1668		5004		11676		25020			
8	3054	13	373,000	PAC-4	2628		7884		18396		39420			

M = Metal, P = Poly, D = Diffusion, S = Substrate, W = Well, X = No connection.

pads for that subarray. As seen in Table I, the PAC-1 was designed to have single probe contact to each subarray. Because of the possibility of probing errors, PAC-2, -3, -4 were designed with double probe pads so that probe contact can be verified. The PAC-1 was designed without a body contact, but such a contact is included on PAC-2, -3, -4 to facilitate the identification of the type of gate-oxide defect.

Experimental Procedures

The arrays were measured in wafer form and in the dark using a computer-controlled parametric test system that has a switch matrix architecture [5]. After a subarray passed a probe-down test, the current through the subarray was measured with 5 V placed across the array, while all other subarrays were unbiased (i.e., electrically floating). For the gate-oxide leakage measurements, the sign of the voltage across a subarray was adjusted to either turn on or turn off the transistor channel. As indicated in Table II, four leakage currents were measured between metal-poly, poly-diffusion (channel on), poly-diffusion (channel off), and poly-body (channel off). For PAC-1, only the first two leakage currents were measured due to the lack of a body contact. For PAC-2, -3, -4, four gate-oxide defects (B, D, E, G) can be identified through the use of the "truth table" listed in Table II. Defects B, D, E are illustrated in Fig. 5. The poly-diffusion current is measured through ammeter ID, and the poly-body current is measured through ammeter IB. When measuring one current, the other ammeter is disconnected.

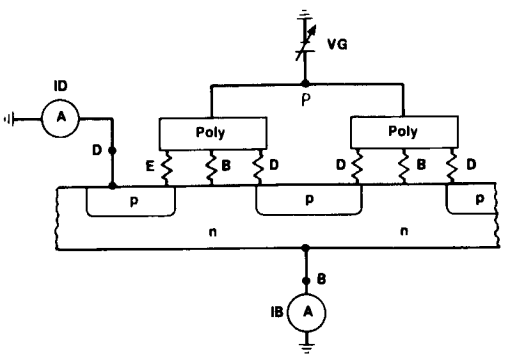


Fig. 5—Cross section of the pinhole array capacitor (without the metal layer), showing poly-bulk (B) shorts, poly-diffusion (D) shorts internal to the array, and poly-diffusion (E) shorts at the edge of the array.

Table II—PAC Measurements and Defect Analysis

Leakage Current Measurements	Channel State	Observation M	N	D	E	B	G
1. Metal-Poly	N/A	S	O	-	-	-	-
2a. Poly-Diff.	On	-	O	S	S	S	S
2b. Poly-Diff.	Off	-	-	O	S	O	S
2c. Poly-Body	Off	-	-	O	O	S	S

Defect
- M = Metal-Poly short
- N = No defect detected
- D = Poly-Diffusion short internal to array
- E = Poly-Diffusion short at edge of array
- B = Poly-Bulk short
- G = Gate-Oxide global short
- S = Short (leakage current > 10 nA)
- O = Open (leakage current < 10 nA).

Measured values for the leakage currents ranged from 0.1 pA to 0.1 mA. The structure is light sensitive and therefore the poly-diffusion and poly-body leakage measurements require a darkened environment. The leakage current values were bimodal; a defect was declared to be detected for leakage currents in excess of 10 nA.

A data set from PAC-1 is illustrated by the wafer maps shown in Fig. 6 for the case of pinholes between metal and n-poly. If the measured current is less than 10 nA, the site is noted by a colon. Otherwise, the site is marked by a plus sign denoting a metal-oxide defect. As seen in the figure, some plus signs are encircled, and these defects are excluded from the data set. These defects are located at Row 5 and Column 7. A photomicrograph of this location, as seen in Fig. 7, reveals that the array was scratched after the poly was patterned and before the metal was deposited. This resulted in a defective deposited oxide layer. Such flaws are

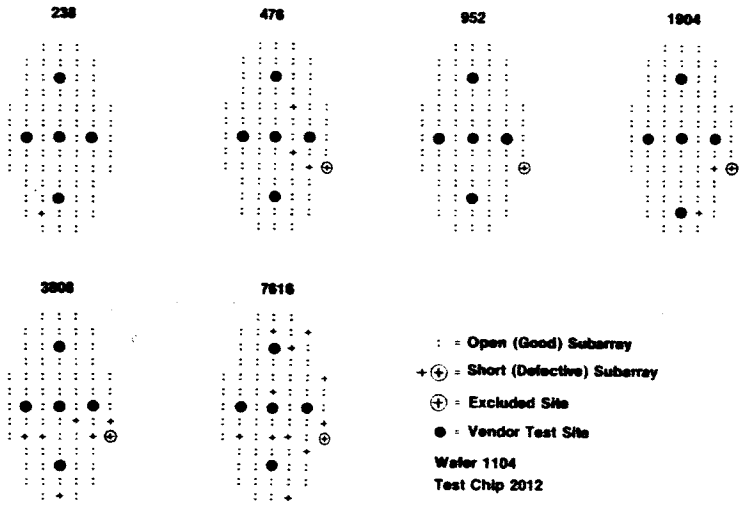


Fig. 6—Wafer maps of the six subarrays found in the pinhole array capacitor, where shorts between the metal and n-poly are denoted by a plus sign (wafer 1104). The number of elements is noted above each wafer map. The number of sites tested for this wafer was 66. (Row number counted from bottom.)

nonlocal global defects and must be excluded from the analysis of local point or oxide pinholes.

Yield Expression

The analysis of point defects is based on the Poisson distribution, where we assume that the occurrence of point defects is random, that the probability of occurrence is small, and that the probability of one defect does not depend on the occurrence of another [11]. The last assumption requires that nonlocal global defects such as scratches and photomask flaws be eliminated from the data set. The expression for the yield of the i'th subarray on a wafer is:

$$Y_i = \exp(-N_i/E) \qquad (1)$$

where N_i is the number of elements per i'th subarray and E is the number of elements per defect. From the method of least squares [12] applied to the above equation, the expression for E is:

$$E = -\sum_{i=1}^{n} N_i^2 \Big/ \sum_{i=1}^{n} N_i \ln Y_i \qquad (2)$$

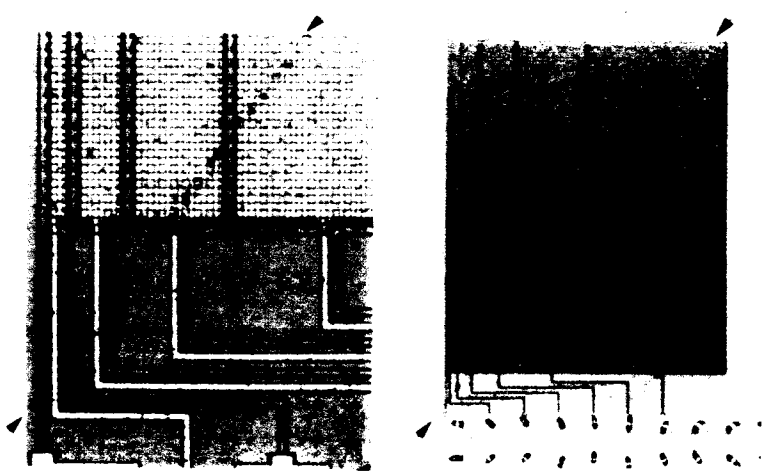


Fig. 7—Photomicrograph of the pinhole array capacitor, showing a nonlocal global defect that has caused a short between the metal and n-poly layers (site Row 5, Column 7, Wafer 1104).

and the standard deviation of E is:

$$E_\sigma = E^2\sigma\Big/\sqrt{\sum_{i=1}^{n} N_i^2} \qquad (3)$$

where the sample variance is:

$$\sigma^2 = \left[\sum_{i=1}^{n}(\ln Y_i)^2 + E^{-1}\sum_{i=1}^{n} N_i \ln Y_i\right]\Big/n \qquad (4)$$

where n is the number of subarrays in the test structure. In the analysis presented in this paper we use the above expression for the variance. Alternatively, we could have replaced n by n-1 for an unbiased estimate of the sample variance [11].

A yield analysis for the metal to n-poly shorts shown in Fig. 6 is presented in Fig. 8. In this figure the i-subscripts for Y and N have been omitted so that Y and N can be thought of as continuous rather than discrete variables. (The yield was calculated using Eq. 5.) Values for E were obtained from the method of least squares given above, where the entire data set was analyzed and where the defective data at location Row 5, Column 7 excluded. Error bars were placed on the graph at the point where $N = E$. The graph illustrates that the value for E is an extrapolation from yield values that are much larger than e^{-1} = 36.8 percent. Unfortunately, this is the situation for PAC included on wafers with working circuits where the area available for diagnostic purposes is very limited.

It is important to estimate the limit of detectability for E. If no defects are detected in all the structures on a wafer, one would like to state that E is greater than some value. Also, a cut-off value for E should be established. For instance, if one defect is found on the wafer, it is not possible to calculate a meaningful value for E. The yield equation can be expressed in terms of the number of defective i'th subarrays per wafer, D_i, or:

$$Y_i = 1 - (D_i/N_s) = \exp(-N_i/E) \qquad (5)$$

where N_s is the number of test sites per wafer. For the case of a very low defect count, where $N_i \ll E$ so that $\exp(-N_i/E) \approx 1-(N_i/E)$, the above equation becomes:

$$D_i = (N_s/E)\,N_i \qquad (6)$$

Employing the method of least squares for $N_i \ll E$:

$$E = N_s\sum_{i=1}^{n} N_i^2\Big/\sum_{i=1}^{n} N_i D_i \qquad (7)$$

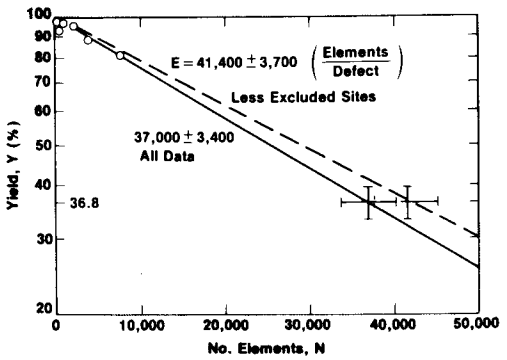


Fig. 8 - Yield curve for the metal to n-poly short data shown in Fig. 6 (Wafer 1104).

The relative error is E is found by differentiating the above equation with respect to D_i or:

$$\delta E/E = -\sum_{i=1}^{n} N_i\delta D_i\Big/\sum_{i=1}^{n} N_i D_i \qquad (8)$$

By combining Eqs. (7) and (8):

$$E = \left(\sum_{i=1}^{n} N_i^2\Big/\sum_{i=1}^{n} N_i\right)(-\delta E/E)/(\delta D/N_s) \qquad (9)$$

where $\delta D_i = \delta D$, which is the error in counting defects in any subarray.

For a maximum allowable relative error in E of -100 percent and for $\delta D = 1$, an estimate for the maximum value of E is:

$$E_{MAX} = N_s\sum_{i=1}^{n} N_i^2\Big/\sum_{i=1}^{n} N_i \qquad (10)$$

For this study, values for *Emax* are listed in Table I, where the values range from near 100,000 to 370,000 elements/defect. For a particular test structure (PAC) and wafer with N_s test sites, *Emax* is used to eliminate excessively large E-values, thus providing a sensible upper bound for E. An estimate for *Emax* can be obtained by considering a test structure composed of one array, N_1. For this case, $Emax = N_s N_1$.

Test Results

Results from Runs 1 and 7 are listed in Tables III and IV, where no data were excluded from the data set. Each Table

Table III—Elements/Defect for Run No. 1*

	Wafer Number						
Region	**1303**	**1304**	**1305**	**1308**	**1310**	**1314**	**Average**
M/NP	31 ± 3	38 ± 4	72 ± 22	34 ± 3	31 ± 3	22 ± 2	38 ± 16
P/ND	183 ± 31	> Emax	288 ± 95	216 ± 53	175 ± 41	272 ± 93	293 ± 178
M/PP	50 ± 8	112 ± 12	54 ± 2	61 ± 8	50 ± 6	34 ± 3	60 ± 24
P/PD	266 ± 50	149 ± 21	145 ± 21	213 ± 48	227 ± 40	155 ± 40	192 ± 46

*Numbers express thousands of elements/defect

lists the E-value and its standard deviation for a number of wafers. Two structures were measured on each wafer, namely a p- and an n-doped PAC. Two measurements were made on each structure. For the n-doped PAC, the metal to n-poly (M/NP) shorts were measured as well as the n-poly to n-diffusion (P/ND) shorts. For the p-doped PAC, the metal to p-poly (M/PP) shorts were measured as well as the p-poly to p-diffusion (P/PD) shorts.

Results in Table III indicate that the poly-to-diffusion gate oxide is less pinhole prone with E-values near 250,000 elements/defect than the metal-poly oxide with E-values near 50,000 elements/defect. Results in Table IV indicate a similar trend as seen in the results in Table III. In Table IV the E-values for the gate oxide are near 100,000 elements/defect and the E-values for the metal-to-poly oxide are near 50,000 elements/defect.

Results from Run 8 are listed in Table V. Notice that there are many entries labeled *Emax*, which indicates that the defect count is very low. There is a relatively large amount of scatter in the values. A similar observation was made by

Table IV—Elements/Defect for Run No. 7*

Region	Wafer Number 7201	7203	7204	7205	Average
M/NP	40 ± 4	38 ± 7	65 ± 9	37 ± 2	45 ± 12
P/ND	57 ± 6	73 ± 9	114 ± 46	105 ± 25	87 ± 22
M/PP	54 ± 6	48 ± 3	36 ± 5	50 ± 6	47 ± 7
P/PD	50 ± 8	78 ± 9	157 ± 37	89 ± 11	101 ± 32

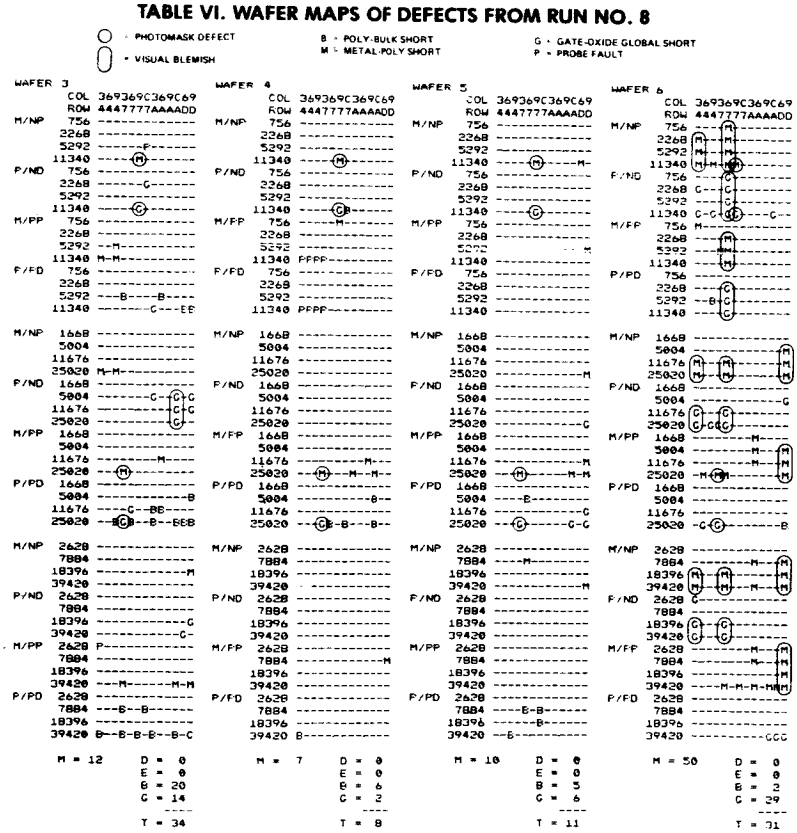
*Numbers in thousands of elements/defect.

Table V—Elements/Defect for Run No. 8*

Wafer	Region	PAC-2	PAC-3	PAC-4
8103	M/NP	> Emax	190 ± 50	> Emax
	P/ND	> Emax	150 ± 100	> Emax
	M/PP	70 ± 7	> Emax	189 ± 49
	P/PD	42 ± 5	35 ± 3	76 ± 18
8104	M/NP	> Emax	> Emax	> Emax
	P/ND	86 ± 22	> Emax	> Emax
	M/PP	> Emax	105 ± 13	> Emax
	P/PD	> Emax	82 ± 19	> Emax
8105	M/NP	86 ± 22	> Emax	> Emax
	P/ND	> Emax	> Emax	> Emax
	M/PP	> Emax	105 ± 13	> Emax
	P/PD	> Emax	100 ± 9	330 ± 185
8106	M/NP	30 ± 4	79 ± 14	95 ± 12
	P/ND	25 ± 4	68 ± 2	199 ± 50
	M/PP	80 ± 31	65 ± 8	71 ± 13
	P/PD	82 ± 34	125 ± 31	190 ± 49

*Numbers in thousands of elements/defect.

TABLE VI. WAFER MAPS OF DEFECTS FROM RUN NO. 8



Ipri [6] about his data. This is attributed to the relatively small number of pinhole defects/wafer and to the occurrence of a number of nonlocal, global defects on some wafers.

An alternative analysis of Run 8 is presented in Table VI, where defects are presented in the form of a linear wafer map. The row and column for each site are given at the top of the table in hexadecimal format. Many of the defects seen in Table VI can be correlated to visual defects seen under a microscope. For example, when three or more adjacent arrays in a structure are defective, the structure is invariably visibly marred. Several defects attributable to photomask flaws are apparent by the appearance of a defect in the same subarray on every wafer. The summary data listed at the bottom of the table indicates that no *D*- or *E*-type defects were located and that the main gate oxide defect is from gate-to-body and not gate-to-diffusion.

Conclusion

A pinhole array capacitor for the characterization of metal-poly and gate oxides has been developed. The analysis involves the determination of the number of elements per defect, E, from yield curves. For this study of 5-μm CMOS-bulk processes, E-values for the metal-poly oxide were typically in excess of 50,000 elements/defect and for gate oxide were typically in excess of 100,000 elements/defect.

By utilizing the switching behavior of the MOS transitors, we were able to distinguish between several kinds of gate oxide defects. Aside from nonlocal global defects, we have reached the tentative conclusion that the gate oxide pinhole defects are gate-to-body shorts rather than gate-to-diffusion shorts. Wafer maps of the defects proved useful in locating defects that occur at the same location on every wafer and hence have a high probability of being due to flaws in the photomasks.

Acknowledgment

The authors gratefully acknowledge the efforts of the MOS Implementation Service of the Information Sciences Institute, University of Southern California, in providing the wafers for Runs 7 and 8. The photomicrographs were supplied by S. F. Suszko, JPL.

References

1. J. Bernard, "The IC Yield Problem: A Tenative Analysis for MOS/SOS Circuits," *IEEE Trans. on ED*, ED-25, pp. 939-944 (1978).
2. K. Saito and E. Arai, "Experimental Analysis and New Modeling of MOS LSI Yield Associated with the Number of Elements," *IEEE J. of Solid State Circuits*, vol. SC-17, pp. 28-33 (1982).
3. C. H. Stapper, "Yield Model for 256K RAMs and Beyond," International Solid State Circuits, pp. 12-13 (1982).
4. C. Timoc, M. Buehler, T. Griswold, C. Pina, F. Scott, L. Hess, "Logical Models of Physical Failures," to be presented at the 1983 Test Conference.
5. M. G. Buehler, T. W. Griswold, C. A. Pina, B. R. Blaes, C. C. Timoc, R. H. Nixon, and S. F. Suszko, "Product Assurance Technology for Procuring Reliable, Custom LSI/VLSI Electronics." JPL Publication No. 83-70, Sept. 1983.
6. A. C. Ipri and J. C. Sarace, "Integrated Circuit Process and Design Rule Evaluation Techniques," *RCA Review*, vol. 38, pp. 323-350 (1977).
7. H. R. Bolin, "Process Defects and Effects on MOSFET Gate Reliability," *Reliability Physics Symposium Proceedings*, vol. 18, pp. 252-254 (1980).
8. M. G. Buehler, T. W. Griswold, C. A. Pina, and C. C. Timoc, "Test Chips for Custom ICs: Six Kinds of Test Structures," *Circuits Manufacturing*, vol. 22, pp. 36-42, June, 1982.
9. T. W. Griswold, "Portable Design Rules for Bulk CMOS," *VLSI Design Magazine*, vol. 3, pp. 62-66, September/October 1982.
10. M. G. Buehler, "Comprehensive Test Patterns with Modular Test Structures: The 2 by N Probe-Pad Array Approach," *Solid State Technology*, vol. 22, pp. 89-94, October 1979.
11. S. L. Meyer, "Data Analysis for Scientists and Engineers," John Wiley and Sons, New York, 1975, pp. 202f.
12. H. D. Young, "Statistical Treatment of Experimental Data," McGraw-Hill Book Co., New York, 1962, pp. 112f.

Martin G. Buehler received a B.S.E.E. and M.S. from Duke University and in 1966 received a Ph.D. in Electrical Engineering from Stanford University. Prior to joining the National Bureau of Standards in 1972, he was a staff member of the Semiconductor Research and Development Laboratories, Texas Instruments, Inc., where he applied electrical measurement techniques to the detection of defects and the profiling of dopants in semiconductor materials. At NBS he led a group in process metrology, designed nearly a dozen semiconductor test structures, was awarded the Department of Commerce's Silver Medal for creative contributions to semiconductor metrology, and initiated a program on self-test techniques for VLSI. Dr. Buehler joined the Jet Propulsion Laboratory in 1981 and is currently the principal investigator of the Product Assurance Technology Program for LSI/VLSI. He is a member of Tau Beta Pi, Eta Kappa Nu, Sigma Xi, the Electromechanical Society, and IEEE.

Thomas W. Griswold earned an A.B. in Physics and in 1953 a Ph.D. in Solid State Physics from the University of California, Berkeley. From 1953 to 1958 he worked on developing germanium and silicon diodes and transistors at the Semiconductor Division, Hughes Aircraft Co. He then moved to Continental Device Corp. as technical director for silicon transistor and IC development until 1972. Dr. Griswold spent the next three years in high-voltage rectifier technology at Semtech Corp. In 1975 he joined the Jet Propulsion Laboratory and supervised the LSI Technology Group. Currently he leads the VLSI Tools and Development Group.

Brent R. Blaes received the B.S. and M.E. degrees in Electrical Engineering from California State Polytechnic University, Pomona, in 1977 and 1979, respectively. He has been with Caltech Jet Propulsion Laboratory since 1977, working on the design and test of high-reliability integrated circuits and test chips.

Cesar Pina received a B.S. in Mathematics from CSLUB in 1968. From 1959 to 1963, he was employed in development and process engineering by Hoffman Electronics Corp., and from 1963 to 1965 he worked there as director of reliability and quality assurance. In 1966 Pina joined Teledyne Semiconductor and later became engineering manager. In 1971 he left Teledyne to start Regulus Semiconductor Corp., where he remained until 1975, when he joined Microsemiconductor Corp. as manufacturing manager. Mr. Pina joined JPL in 1980 and currently leads the LSI Technology Group.

2.3.2 Contact-Resistance Process Cliff[1]

2.3.2.1 Discussion. Several approaches have been investigated over the past two years to obtain a measure of the quality of the contact resistance between metal and a diffused or polysilicon layer. These approaches have included the use of both short and very long contact strings as well as arrays of contacts with different sizes to determine the contact resistance process cliff.

The results of the evaluations of the contact strings have been particularly disappointing. Initially we assumed that the contact strings could be used to detect the density of open contacts and provide fault density statistics similar to those we have obtained with the capacitor arrays described elsewhere in this report. However, our studies with large strings having as many as 120,000 contacts in a single string seldom revealed an open contact. We next thought that the strings might be useful in detecting marginal contacts, that is, in detecting contacts whose resistance is a few standard deviations from the mean contact resistance. But this approach was discarded using the following reasoning: If a string has 1000 contacts and each contact has a resistance of 1 ohm, then the resistance of the string will be 1000 ohms. If the string has an anomalous contact with a resistance of, say, 10 ohms, its presence in the string will go undetected, for it represents only a 1% change in the value of the bad string over the value of a good string. Implicit in the above argument is the assumption that there are enough variables in the fabrication of the string that one can not expect to fabricate a contact string with a 1% tolerance. The variables other than the contact resistance that contribute to this tolerance are the line-width and sheet resistance of the semiconducting material (diffusion or polycrystalline silicon).

A different approach to assessing the quality of metal-diffusion or metal-polysilicon contacts consists of using contact resistors of various sizes and determining the relationship between the contact resistance and the contact area. This relationship has been found to be process and/or manufacturer dependent and provides a number, the "Contact Process Cliff," that serves as a measure of the quality of the contacts of a particular lot or wafer. This approach, of course, also fails to provide statistics on the failure of individual contacts. It does, however, enable the evaluator of the lot or wafer to make a statement about the overall quality of the lot or wafer.

Essentially, the method consists of fabricating a number of square contacts of different sizes, typically differing from each other by 10% or less in the linear dimensions, and decreasing from 10% above the design rules to a size equal to 50% smaller than the design rule. The resistors are then measured at a constant current density, and the resulting contact resistance is plotted against the as-drawn contact area. The contact area for which any further decrease in the area of the contact results in a large jump in the value of the contact resistance is designated as the "Contact-Resistance Process Cliff." The cliff, therefore, denotes the minimum value of the as-drawn contact size that produces contacts with "reasonable" contact resistance values. That is, the "Process Cliff" is a measure of the tolerance

[1]This section was prepared by C. A. Pina.

in the width of the contact design rule. It indicates the variability in the width a given process may have and still produce parts with acceptable values of contact resistance.

Results from metal-n-diffusion contact resistors are shown in Table 2.3.2-1. These results were obtained from four-terminal Kelvin contact resistors [1] designed with different square contact areas. The current density was maintained at 40 $\mu A/\mu m^2$ during the measurement. As-drawn dimensions were used to establish the contact area.

In the table the contact width and contact resistance measured for contacts slightly smaller than the cliff are denoted as Wc1 and Rc1, respectively. The contact width and contact resistance measured for contacts slightly larger than the cliff are denoted as Wc2 and Rc2, respectively. The width of the contact at the design rule is 3 μm and the contact resistance at the design rule is noted as Rcdr. Referring to Figure 2.3.2-1, Wc1 is the contact size at the "top" of the cliff, and Rc1 is the coresponding contact resistance. Similarly, Wc2 is the contact size at the "bottom" of the cliff, with Rc2 as the corresponding value of contact resistance. Rcdr is the contact resistance value at the design rule limit, Wcdr, of 3-μm square contact openings.

Only the metal-n-diffusion contact resistance results are summarized in the table; however, contact resistance measurements for metal-p-diffusion and metal-poly exhibited similar behavior, as seen in Figures 2.3.2-2 through -4.

The data in the table shows that the contact resistance values from runs B, C, and G were excellent. That is, the contact resistance Rc2 is low for a contact width Wc2 that is considerably smaller than the design rule width of 3 μm. Runs D and E are judged to be marginal since the Wc2 value equals the design rule width of 3 μm and the contact resistances are high. Runs A and F may be satisfactory depending on the process specifications.

The runs shown in Table 2.3.2-1 were produced by three different manufacturers, and the relative quality of the manufacturers can be readily determined from the data. Since manufacturing processes are not perfect and process variations about the design rule limit must be expected, the data shows that Manufacturer I has the best contact resistance values. Thus the contact-resistance process cliff approach to evaluating contact resistance should prove to be a valuable tool.

2.3.2.2 Reference.

1. S. J. Proctor and L. W. Linholm, "Direct Measurement of the Interfacial Contact Resistance, End Contact Resistance, and Interfacial Contact Layer Uniformity," IEEE Trans. Electron Devices, ED-30, 1535-1542 (1983).

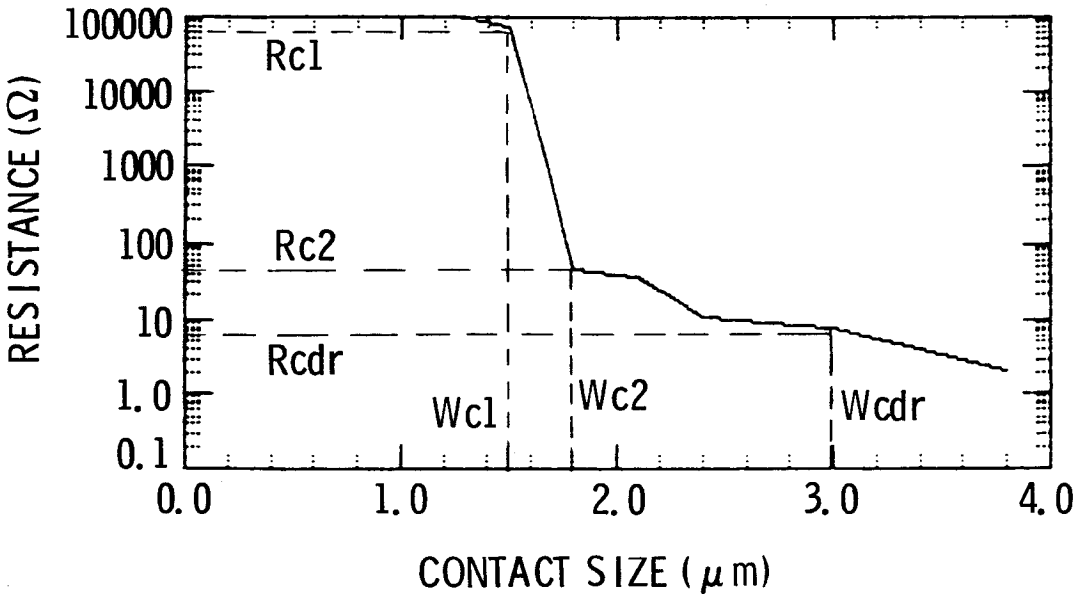


Figure 2.3.2-1. Metal-n-diffusion contact-resistance process cliff

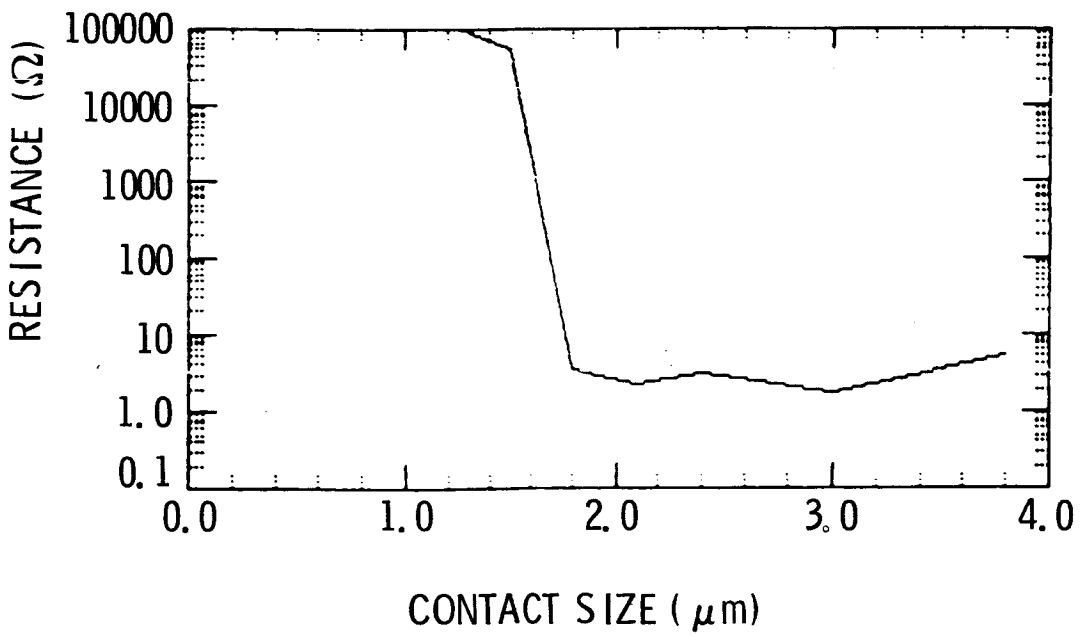


Figure 2.3.2-2. Metal-p-diffusion contact-resistance process cliff

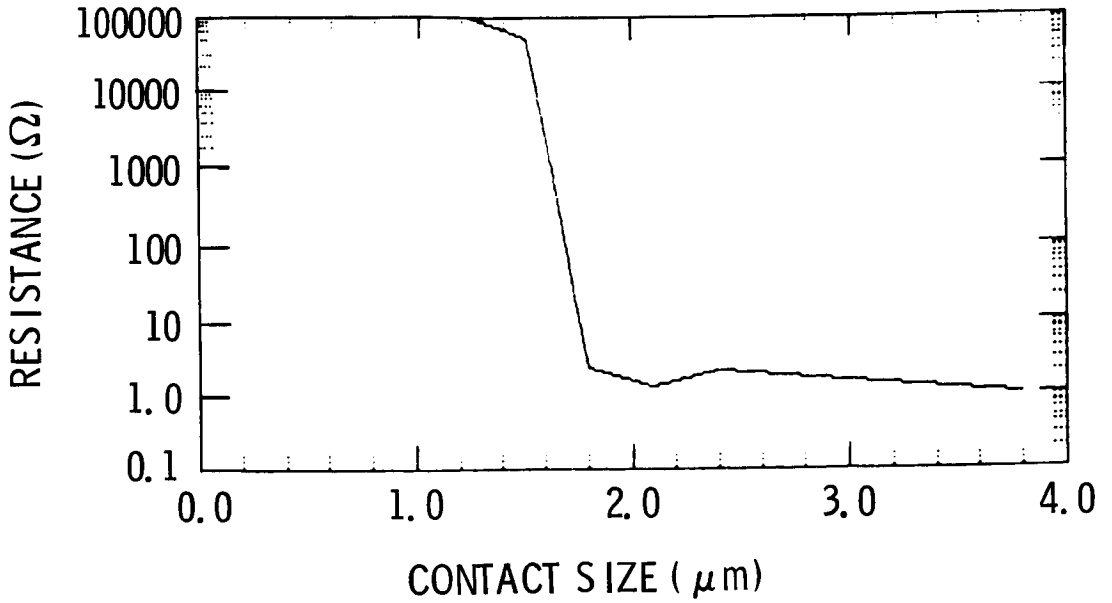


Figure 2.3.2-3. Metal-n-poly contact-resistance process cliff

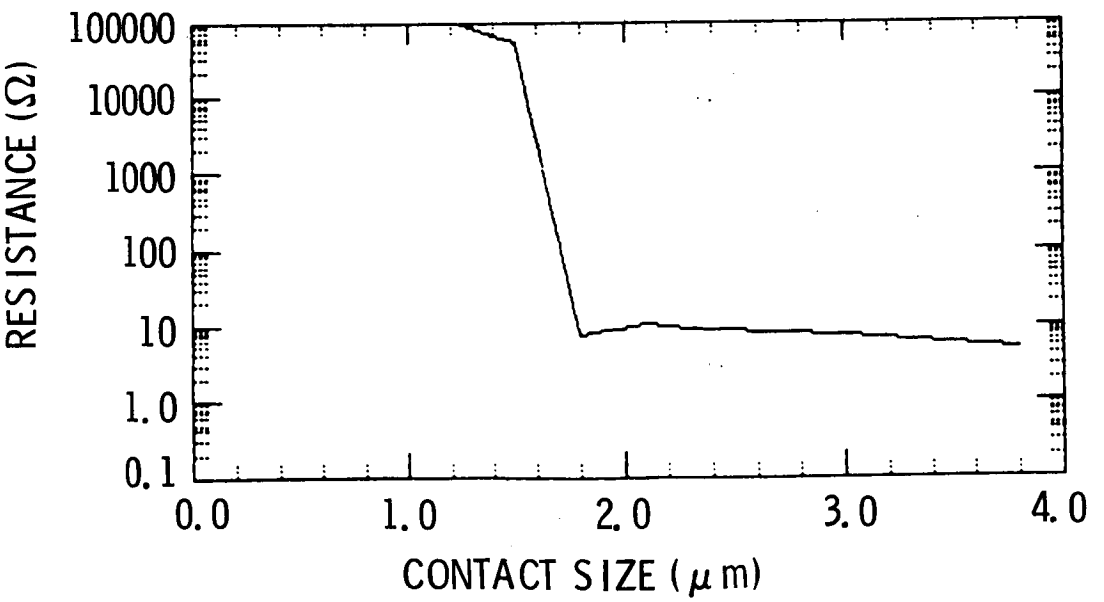


Figure 2.3.2-4. Metal-p-poly contact-resistance process cliff

Table 2.3.2-1. Metal-n-Diffusion Contact-Resistance-Process-Cliff Results

RUN	MANUFACTURER	CONTACT RESISTANCE PROCESS CLIFF				CONTACT RESISTANCE DESIGN RULE
		Wc1 (μm)	Rc1 (ohm)	Wc2 (μm)	Rc2 (ohm)	Rcdr (ohm)
A	I	1.5	60k	1.8	50.0	8.0
B	I	1.5	5k	1.8	5.0	3.0
C	I	1.5	9k	1.8	2.0	2.0
D	II	2.4	20k	3.0	20.0	20.0
E	II	2.4	35k	3.0	100.0	100.0
F	II	2.4	30k	2.7	10.0	8.0
G	III	-*	-*	2.1	8.0	4.0

*The smallest contact width submitted on this run was 2.1 μm.

2.3.3 Addressable Inverter Matrix[1]

2.3.3.1 Discussion. Since publishing the data in [1], we have decreased the data acquisition time for the inverter matrix by a factor of 10. This was accomplished by changing the test system operating software.

The testing of the matrix can take two forms. The inverter transfer curve can be analyzed using a 5-point analysis consisting of VLOW, VHIGH, VINV - 25 mV, VINV, and VINV + 25 mV, where VINV is measured with the inverter output connected to its input, VLOW is the inverter output voltage for the input high, and VHIGH is the inverter output voltage for the input low. From these values the inverter GAIN is determined from:

$$\text{MAGNITUDE GAIN} = [(\text{VINV} - 25\ \text{mV}) - (\text{VINV} + 25\ \text{mV})]/50\ \text{mV}.$$

Alternatively, the transfer curve can be analyzed using a 51-point method in which the input voltage is stepped on 100-mV increments and the output voltage is measured. The voltmeter used in the above measurements was a Hewlett-Packard 3456A which has a 10-μV resolution on the 10-V scale and a maximum data acquisition rate of 0.02 seconds per voltage reading. The overall data acquisition system is described elsewhere [2].

The 5-point transfer curve analysis involves clearing the shift register, selecting a column, measuring VINV along the column, measuring GAIN, VLOW, and VHIGH at each inverter in a column, and then selecting the next column to be analyzed. This sequence was chosen to minimize the switching of the tester's mechanical switch matrix. The computer (LSI-11) was initially programmed using the software tool CRUNCH, and test times were about 32 minutes to acquire data from the 222 inverters in the matrix. This required the measurement of 1110 voltage points, and the data acquisition rate was 1.7 seconds per voltage reading. The computer was then programmed directly in Fortran and the data acquisition time was reduced to 2.6 minutes. The data acquisition rate was 0.14 seconds per voltage reading, which includes the time to operate the shift register and the mechanical switch matrix. The reason for the dramatic reduction in the acquisition time is due to the conversion of the test program from CRUNCH to Fortran. CRUNCH is a line interpreter with a very slow execution time.

In some cases a "complete" inverter transfer curve is desired rather than the 5-point transfer curve. For this analysis the input voltage to the inverter is stepped in 100-mV increments, which requires the measurement of 51 data points per inverter. The total number of data points per matrix is 11,322 voltage readings. In this case, the time to measure the transfer curve of the 222 inverters in the matrix using the direct Fortran programming is 20 minutes. The data acquisition rate is then 0.09 seconds per voltage reading, which includes the time to operate the shift register and mechanical switch matrix, and the plotting time.

[1]This section was prepared by H. R. Sayah and M. G. Buehler.

In conclusion, this study has shown that the software used to program the data acquisition instruments can have a major effect on the data acquisition time. In this case the data acquisition time was reduced by an order of magnitude by programming directly in Fortran rather than using a line-interpreter-based program.

2.3.3.2 References.

1. M. G. Buehler and H. R. Sayah, "Addressable Inverter Matrix for Process and Device Characterization," Solid State Technol., 28, 185-191 (May 1985).

Note: This article is reprinted in its entirety as it appeared in the journal, at the end of this section.

2. M. G. Buehler, T. W. Griswold, C. A. Pina, B. R. Blaes, C. C. Timoc, R. H. Nixon, and S. F. Suszko, "Product Assurance Technology for Procuring Custom LSI/VLSI Electronics," JPL Publication 83-70, Pasadena, California (September 1983).

Addressable Inverter Matrix for Process and Device Characterization

Martin G. Buehler Hoshyar R. Sayah
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

The addressable inverter matrix consists of 222 inverters each accessible with the aid of a shift register. The structure has proven useful in characterizing the variability of inverter transfer curves and in diagnosing processing faults. For good 3 μm CMOS bulk inverters investigated in this study, % standard deviation of the inverter threshold voltage was less than one percent and the inverter gain (the slope of the inverter transfer curve at the inverter threshold voltage) was less than three percent. The average noise margin for the inverters was near 2 volts for a V_{DD} of 5 volts. The specific faults studied included undersize pull-down transistor widths and various open contacts in the matrix.

USING MICROELECTRONIC TEST STRUCTURES for process characterization is now common. Yet these structures often consume an appreciable portion of the area of the wafers and are considered an automatic yield loss. The area issue is of long standing and has tended to limit the use of test chips. In more recent times, the shift from whole-wafer lithography to direct step-on-wafer lithography has placed an even further emphasis on the efficient utilization of wafer area for diagnostic purposes.

Test chips can be "pad intensive" where each element in a test structure (such as a transistor) is connected to a separate probe pad. This approach to test chip layout has the advantage of eliminating the interference of one structure on another. The disadvantage to this approach is that only a limited number of structures can be sampled so that it may be difficult to establish a meaningful characterization of device parameters and process faults.

For these reasons it was found desirable to undertake the development of a set of addressable test structures that use addressing schemes to access individual elements in a matrix. The results for an addressable inverter matrix that allows the characterization of the dc transfer curves of 222 inverters are discussed.

Addressable test structures have their origin in such circuits as random access memories and programmable logic arrays. One recently developed test structure with addressing circuitry is a photomask misalignment structure that uses a shift register to address various elements [1]. Matrix concepts have been used to access clusters of transistors where the rows and columns are connected to individual probe pads [2]. In this transistor matrix, the addressing is performed off-chip by a parametric tester's mechanical switch matrix connected through the probes to the test structure. In such a structure the number of elements in the matrix is limited by the number of probe points. For 20 probe points, the number of elements is limited to 100.

In the matrix approach described here, the number of elements (inverters) does not depend on the number of probe points. Also the inverter matrix can be used to provide a complete characterization of each inverter transfer curve. Basically, this structure can be thought of as providing a digital assist to analog measurements. The disadvantage in using an addressable test structure, however, is found to be the difficulty in separating faults that appear in the ancillary circuitry from faults that occur in the inverters in the matrix. To aid in diagnosing the ancillary circuitry, reference sites were added to the matrix.

The goal of this effort is to explore the effectiveness of the addressable inverter matrix in analyzing the dc characteristics of CMOS inverters and in providing diagnostic information. It is shown that the inverter matrix is useful in: (a) characterizing the variability of the inverter transfer curves; (b) evaluating the effect of undersize pull-down inverter-transfer widths on transfer curves; and (c) pinpointing the location of open contacts in inverter-transmission gate cells.

Addressable Inverter Matrix

The test structure, shown in Fig. 1, consists of a shift register located on the left side of the structure and a matrix of inverter-transmission gate cells connected to a 2 by 10 probe array at the bottom of the test structure [3]. The structure is confined to a square area that is 1.6 mm on a side. This layout allows for combining of the test structure with other structures into a test chip [4] that permits probing of all the structures with the same probe array. Details of the matrix are shown in Fig. 2 where it is seen that all the inverter inputs are connected to the structure input labeled INV_{IN}. The inverter outputs are connected to row-addressable transmission gates. The transmission gates are connected to column busses that terminate at the probe pads labeled INV_{OUT}. The shift register is used to turn on a row of transmission gates by providing a proper bias to Q and $\overline{Q}$. Power supply volt-

Reprinted with permission of Solid State Technology, published by Technical Publishing, a company of Dun & Bradstreet.

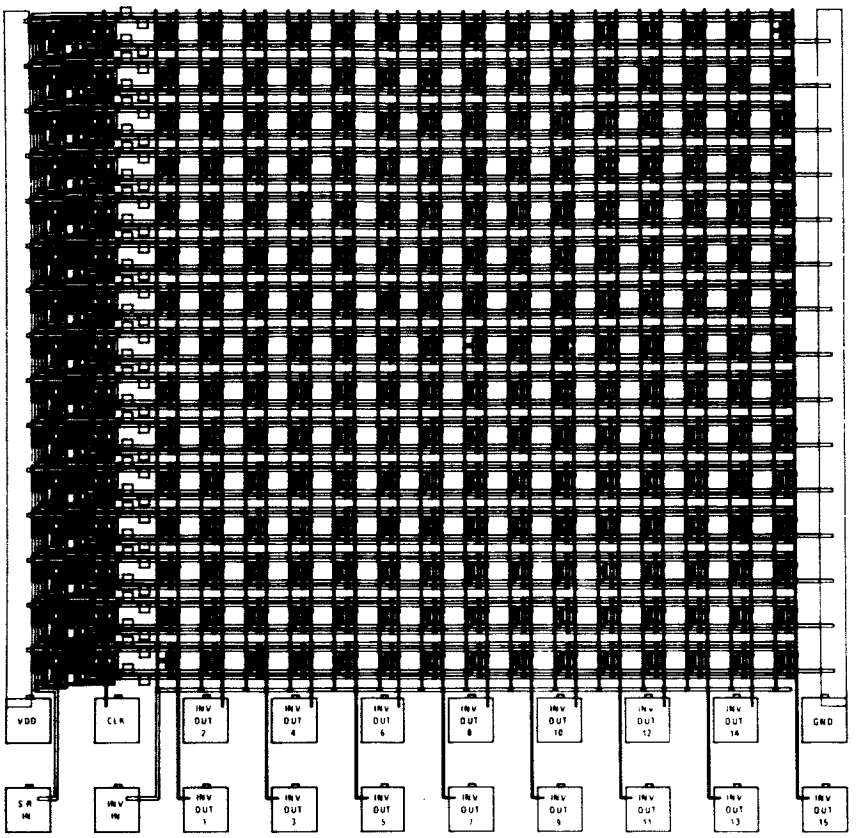


Fig. 1—Addressable inverter matrix designed to be probed by a 2 by 10 probe pad array.

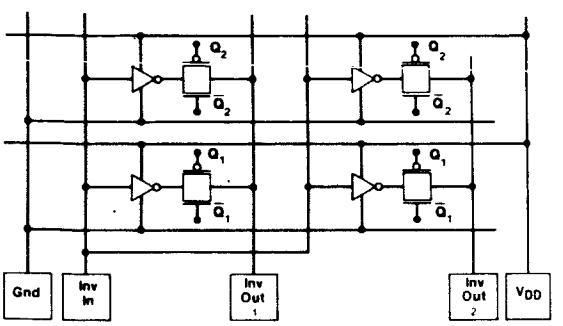


Fig. 2—Schematic layout of four inverter-transmission gate cells.

age V_{DD} and the ground (GND) busses are shown explicitly in Fig. 2 to illustrate that each inverter has the same low series resistance in its power lines. This conventional static CMOS shift register uses inverters and transmission gates in a master-slave D-flip flop configuration [5].

In order to verify the operation of the shift register, three reference sites were introduced at the lower left, middle, and upper right of the matrix. A reference site can be seen in Fig. 3. In these sites the inverters were removed and the input to the transmission gate connected to the structure input. When measurements are made at these sites, the output voltage is expected to equal the input voltage for good devices. Thus, these sites serve as markers and are used to check that the shift register has advanced to rows 1, 8, and 15. The transfer curves for these reference sites are straight lines where $V_{OUT} = V_{IN}$ and they can be seen as diagonal lines in Fig. 4.

The structure was fabricated in a 3-μm CMOS bulk p-well self-aligned poly-gate isoplanar process. The inverters were designed with gate features of $W_p/L_p = 6.0/3.0$ for the p-channel pull-up transistor and $W_n/L_n = 4.5/3.0$ for the n-channel pull-down transistors where the dimensions are in micrometers.

Test Equipment and Test Program

The test equipment used in this study consists of the following instruments: a word generator for pulsing the shift register; a 5 volt power supply; a digital-to-analog converter with 0.5 millivolt resolution for generating inverter input voltages; and a digital voltmeter with 10 gigaohms input impedance and 100 microvolt resolution for measuring inverter outputs. These instruments are computer controlled, and a test program was developed to test the matrix which consists of 15 rows and 15 columns of inverter-transmission gate cells. The shift register is used to select the row to be measured. However, before this can be done the shift register must be cleared since it is in an unknown state when

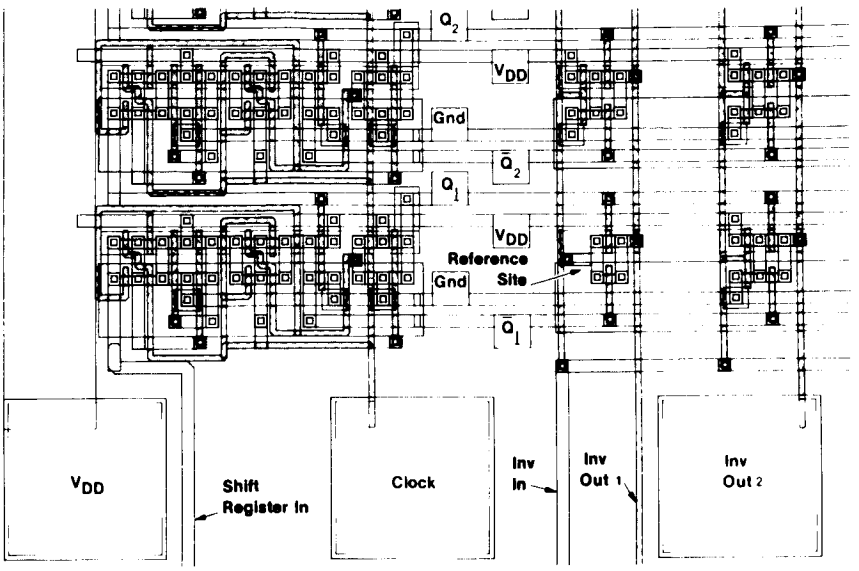


Fig. 3—Detailed layout of the addressable inverter matrix showing the reference site in the lower left hand corner of the matrix. The polycrystalline layer is shown in a gray tone.

power is first applied. After clearing, a row is selected by clocking a pulse through the shift register to select one of the 15 rows.

In order to avoid taking faulty data, the program monitors the power supply voltage before taking each data point. If latch-up or a short occurs, the program will flag the data if the power supply voltage drops below 4.75 volts. The power supply is current limited to about 20 mA to protect the probe card against excessively large current surges.

The structure was measured in two modes. When complete transfer curves were desired, the input voltage was stepped in increments of 0.1 volts and the output voltage recorded. In order to characterize inverters quickly, a five-voltage measurement was performed. From the inverter output voltages we determined V_{HIGH} (V_H), V_{LOW} (V_L), V_{INV}, and Gain. V_t was measured at the inverter output for a high input. V_H was measured at the inverter output for a low input. V_{INV} was determined by connecting the output of the inverter to its input and measuring the resulting voltage, termed the inverter threshold voltage. The Gain is the magnitude of the slope of the inverter transfer curve in the vicinity of V_{INV}. The Gain was determined from two inverter output voltage measurements. In one measurement the inverter input is forced to 25 millivolts less than V_{INV} and in the other the voltage is forced to 25 millivolts greater than V_{INV}. The Gain is the magnitude of the difference between the resulting voltages divided by 50 millivolts.

The data was taken in a sequence which was designed to minimize the switching of the tester's mechanical switch matrix. First V_{INV} is measured at each cell along a specific column. Then Gain, V_t and V_H are measured for each cell in a column. The sequence is not important for good cells but becomes important for faulty cells. For faulty cells the results can be test sequence and time dependent because these cells can have transistors that may be stuck on, stuck off, or have high leakage current.

Table I—Test Results from Two Good Inverter Matrices

Parameter (Units)	Run A Location 3, 4 Mean	Run A % SD*	Run B Location 3, 5 Mean	Run B % SD
V_H (volts)	5.00	1E-4	5.00	1E-4
V_t (volts)	7E-6	137	1E-5	369
V_{INV} (volts)	1.96	0.58	2.19	0.71
Gain	22.60	2.24	12.90	2.74
N (volts)	1.87	0.57	2.02	0.72

*Standard Deviation

Results

In this Section we discuss results from: (a) matrices with good inverters; (b) matrices with undersize pull-down transistor widths; and (c) a matrix with open contacts.

Results from Good Inverter Matrices

Results from two good inverter matrices are listed in Table I. From the V_H and V_t values, it is seen that the inverters are pulling to V_{DD} and GND respectively. It is also seen that the V_{INV} and Gain values are tightly distributed. Note that the % standard deviation for the V_t values are habitually very large because the calculation requires the standard deviation to be divided by the mean which in this case is very close to zero. As seen in the table, the % standard deviation for V_{INV} is less than one percent and for Gain is less than 3 percent. By combining V_{INV} and Gain, the noise margin N was calculated from the following equation which was derived using the maximum square approach [6] and piece-wise linear approximation to the transfer curve [7]:

$$N = (1 - 1/\text{Gain})\, V_{INV} \qquad (1)$$

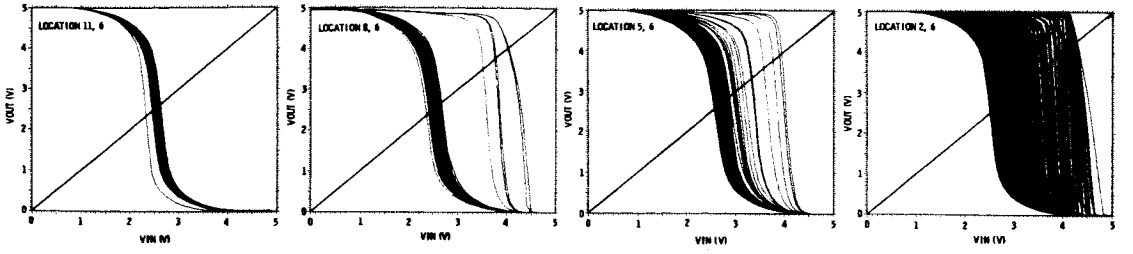


Fig. 4—Inverter transfer curves taken from four addressable inverter matrices located at four sites on a wafer. These results are due to undersize pull-down transistor widths, (Run C).

Table II—Test Results from Four Faulty Inverter Matrices

	Run C							
Parameter (Units)	**Location 2, 6**		**Location 5, 6**		**Location 8, 6**		**Location 11, 6**	
	Mean	**% SD***	**Mean**	**% SD**	**Mean**	**% SD**	**Mean**	**% SD**
V_H (volts)	5.00	6E-3	5.00	2.1E-2	4.96	0.17	5.00	8E-3
V_t (volts)	0.37	292	1E-5	186	0.02	1051	2E-3	325
V_{INV} (volts)	3.11	-17.21	2.69	6.68	2.54	7.89	2.55	1.15
Gain	19.70	79.22	7.95	13.00	9.27	7.16	11.70	3.68
N (volts)	1.79	28.62	2.02	8.01	2.19	8.18	2.24	1.23

*Standard Deviation

The noise margin values listed in Table I are well above the worst case noise margin of 25 percent of V_{DD} or 1.25 volts [8].

Results from Matrices with Undersize Pull-Down Transistor Widths

Results from four inverter matrices are listed in Table II. These results were obtained at four locations across the diameter of a 4 inch diameter wafer. The wafer had 13 chips across its diameter. It is apparent from the large % standard deviations in V_{INV} and Gain that the inverter transfer curves are not as tightly distributed as those listed in Table I. A closer examination of the data reveals that the percent standard deviation in V_{INV}, Gain, and Noise Margin generally decreases moving from location 2,6 to 11,6. A further clarification is shown in Fig. 4 where the transfer curves for each location are shown. The location of the outlier curves was found to be randomly distributed throughout the matrix.

This behavior can be explained by variations in the width of the pull-down transistor. This width was designed to be fabricated at the minimum layout rule width. After fabrication the diffusion width was much smaller than the minimum layout rule as determined by split-cross-bridge resistor measurements [4]. Instead of being at the layout rule width of 4.5 μm, the average width for this wafer was 1.16 ± 0.10 μm. The trends seen in Fig. 4 can be explained by the dependence of V_{INV} on W_n. The expression for the inverter threshold voltage [8], which is derived in the Appendix, is:

$$V_{INV} = \frac{V_{DD} + \sqrt{B_o}V_{tn} - |V_{tp}|}{1 + \sqrt{B_o}} \qquad (2)$$

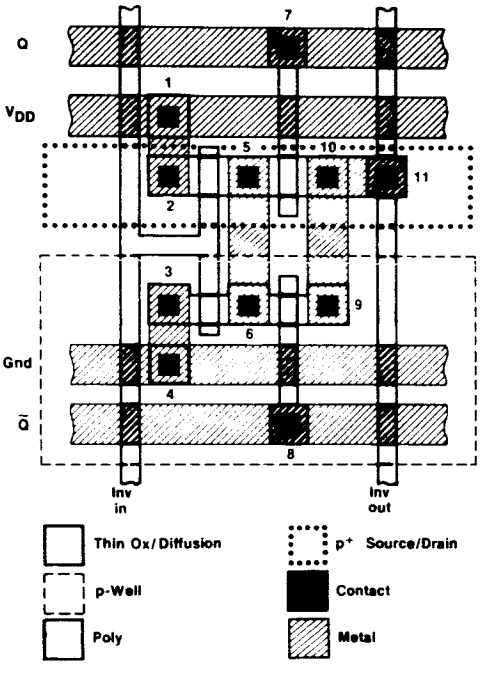


Fig. 5—Layout of the inverter-transmission gate cell showing the location of the 11 contacts.

where V_{tn} is the n-channel threshold voltage, $|V_{tp}|$ is the magnitude of the p-channel transistor threshold voltage, and

$$B_o = (K_n W_n/L_n)/(K_p W_p/L_p) \qquad (3)$$

where K_n and K_p are the n- and p-channel transistor conduction factors (sometimes called KPRIME), W_n and W_p are the n- and p-transistor gate widths, and L_n and L_p are the n- and p-channel transistor gate lengths. In the limit as W_n goes to zero, the inverter threshold voltage becomes:

$$V_{INV} = V_{DD} - |V_{Tp}| \quad (4)$$

For the inverters used in this study, $|V_{Tp}|$ was 0.98 ± 0.02 volts and V_{Tn} was 0.98 ± 0.10 volts. For a V_{DD} of 5 volts, V_{INV} can be expected to reach 4 volts and this is the trend observed in Fig. 4, especially for location 2,6. Note that $|V_{Tp}|$ is approximately given by the point where the transfer curve just rises above the $V_{OUT} = 0$ line.

Results from a Matrix with Open Contacts

The occurrence of open contacts was observed using the addressable inverter matrix. Open contacts can occur at any or all of the 11 contacts in the inverter-transmission gate cell as seen in Fig. 5.

Photomicrographs of selected cells are shown in Fig. 6 where the cell at column 3, row 8, is unfaulted. As seen in the figure, the cell at 4,8 is missing the #3, #4, #6, and #8 contacts (also see Fig. 5); the cell at 3,7 is missing the #4, and #8 contacts; the cell at 4,7 has all contacts open.

A list of the open contacts for the matrix under study is given in Fig. 7. This list was determined from photomicrographs of each cell similar to those shown in Fig. 6. The corresponding inverter parameters for the cells shown in Fig. 7 are given in Fig. 8, where faulty V_L and V_H values are circled. For this study a fault was defined as V_L greater than 0.0 volts and V_H less than 5.0 volts. The faults observed in this study either prevented the cell from pulling up to V_{DD} or down to GND or both.

An analysis of the inverter-transmission gate cell with singly occurring open contacts is given in Table III. In the table, open contacts at #1 (substrate) and #4 (well) were omitted because other cells in a given row are connected to these layers. Thus, the likelihood of having no contact to these layers in a given row was considered remote. When certain contacts are open, a cell's apparent output will be determined by the previous state of the output bus. For such cases the state of the output bus is denoted $Q-$.

As seen in Table III, an open contact at contacts #2, #3, and #11, can lead to a $Q-$ state. The value of $Q-$ is determined by the test sequence. In this study the test sequence begins with the measurement of V_{INV} at each cell in a column. Then, the Gain, V_L and V_H are measured at each cell starting at the bottom of the column and progressing to the top of the column. Before testing the next cell up the column, the input to the cell under test is biased to a high state, which leaves the output bus in a low state.

The results shown in Fig. 8, Column 1, can be explained by realizing that since the good cell at 1,4 leaves the output bus in a low state, the cells at 1,5 1,6 and 1,7 which have an open #11 contact, are unable to pull the output bus high. A similar situation holds in Column 5 for cells 5,5 and 5,6. The cell at 5,7 is connected to the output bus but an open #2 con-

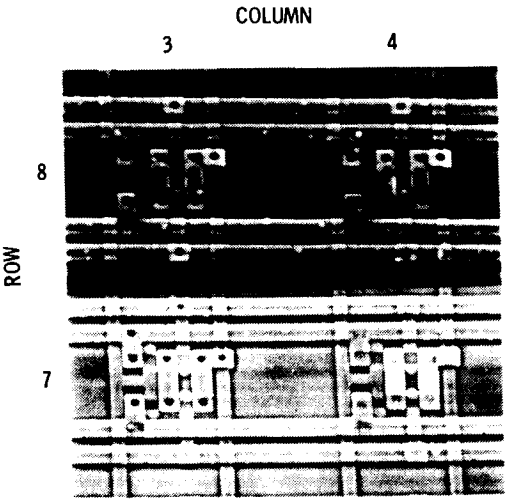


Fig. 6—Photomicrographs of four inverter-transmission gate cells with open contacts. When a contact is present it appears as a dark spot, (Run D).

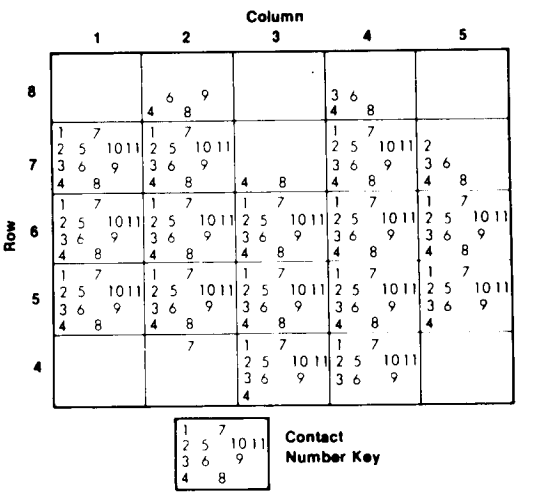


Fig. 7—The location of open contacts in a portion of the addressable inverter matrix. The numbers refer to the contacts labeled in Figure 5, (Run D).

Table III—Inverter-Transmission Gate Output for Open Contacts

Open at Contact Number	Output V_L	Output V_H
None	0	V_{DD}
2	0	Q-
3	Q-	V_{DD}
5 or 6	0	V_{DD}
7(ON)	0	V_{DD}
7(OFF) and/or 10	0	$V_{DD}-\|V_{Tp}\|$
8(ON)	0	V_{DD}
8(OFF) and/or 9	V_{Tn}	V_{DD}
11	Q-	Q-

tact prevents this cell from pulling the output high and so the output remains low.

If the #7 or #8 contact is open, the transmission gate will not be fully functional. These open contacts lead to floating gate transistors which are modeled as being in a conducting (ON) state or a non-conducting (OFF) state. Our data is best explained if we assume that for an open #7 contact, the p-channel transistor is stuck on and for an open #8 contact, the n-channel transistor is stuck off. An open #7 contact occurred at cell 2,4. Under the assumption that the p-channel transistor in the transmission gate is stuck on, the open #7 contact is undetected in evaluating cell 2,4. An open #8 contact occurred at cell 3,7. The assumption that the n-channel transistor in the transmission gate is stuck off explains the observation that this cell only pulls down to within V_{Tn} of GND.

The results for cells 3,4 3,5 3,6 4,4 4,5 4,6 and 4,7 are nearly identical. They can be explained by the leakage of current onto the output bus through leaky p-channel transistors in those transmission gates found in good cells connected to the output bus. This result was confirmed by closely examining the V_L and V_H values. These values slowly drifted upwards as measurements were taken up the column indicating the leakage of current onto the output bus. The results for cells at 2,5 2,6 2,7 2,8 and 4,8 are difficult to explain and remain under study.

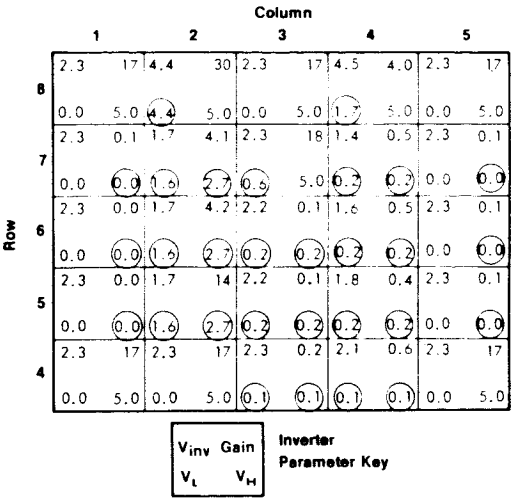


Fig. 8—Inverter parameters for the same portion of the addressable inverter matrix shown in Figure 7. Faulty V_L and V_H values are circled, (Run D).

Ancillary Circuitry

At the outset it was mentioned that the ancillary circuitry must also be analyzed for faults. In this case faults in the shift register and the transmission gates must be identified. The three reference sites are used to verify that the shift register has advanced to rows 1, 8, and 15. The transfer curves for the transmission gates at these sites can be seen in Fig. 4 as the 45 degree equality line. Metal shorts in the shift register that prevent the propagation of the signal through the shift register were observed. Faulty transmission gates have been observed and these give rise to low Gain values for the inverter-transmission gate cell.

Conclusion

The test matrix used in this study contained 222 3-μm CMOS bulk inverters and was fabricated in a square area approximately 1.6 mm on a side. The addressable inverter matrix has proved to be an effective tool in both analyzing the dc characteristics of inverters and in identifying the several types of faults. In this study we were able to characterize the effect of undersize n-channel transistor widths on the inverter transfer curves. This effect causes the transfer curves to shift to higher V_{INV} values that approach $V_{DD}-|V_{Tp}|$. The effect of various open contacts in the inverter-transmission gate cell were observed. Some open contacts (#2, #3, and #11) produce results that depend on results from previously measured cells. Other open contacts [#5 and/or #6, #7(ON), #8(ON)] can go undetected provided leakage currents are low.

Acknowledgments

The authors wish to acknowledge the effort of Robert Nixon in designing the structure and to thank Cesar Pina for supplying the parametric test equipment, Brent Blaes for the inverter data in figure 4, and Ronald Ruiz for the photomicrographs. They also thank Terry Walker of Stanford University for suggesting the shift register design. Also, the efforts of the MOS Implementation Service, Information Sciences Institute, University of Southern California in fabricating the test structures are greatly appreciated. This work was sponsored jointly by the National Aeronautics and Space Administration, the Defense Advanced Research Projects Agency, and the National Security Agency under contract number NAS-918.

Appendix

The inverter threshold voltage, expressed by Eq. 2, was derived for a CMOS bulk inverter with enhancement mode transistors operating in the saturation region. In this region the current through the n-channel pull-down transistor is given by:

$$I_n = \beta_n(V_i-V_{Tn})^2/(1-\lambda_n V_o) \quad \text{(A1)}$$

and the current through the p-channel pull-up transistor is given by:

$$I_p = \beta_p(V_{DD}-V_i-|V_{Tp}|)^2/(1-\lambda_p(V_{DD}-V_o)) \quad \text{(A2)}$$

where the inverter input voltage is V_i, the output voltage is V_o, and λ_n and λ_p are the n- and p-channel length modulation factors. By including channel length modulation in the model, the Gain has a finite value at V_{INV} for λ_n and λ_p non zero. In addition

$$\beta_n = K_n W_n/L_n \quad \text{(A3)}$$

and

$$\beta_p = K_p W_p/L_p \quad \text{(A4)}$$

By setting $I_n = I_p$ and $V_o = V_i = V_{INV}$, taking the square root of the resulting equation, rearranging terms, the result is Eq. 2 where B appears in place of B_o:

$$B = \frac{\beta_n(1 - \lambda_p(V_{DD} - V_{INV}))}{\beta_p(1 - \lambda_n V_{INV})} \tag{A5}$$

In the limit where $\lambda_n = \lambda_p = 0$, $B = B_o$.

References

1. B. M. M. Henderson, A. M. Gundlach, A. J. Walton, "Integrated-Circuit Test Structure which Uses a Vernier to Electrically Measure Mask Alignment" *Electronics Letters*, vol. 19, pp. 868-869 (1983).
2. M. G. Buehler, L. W. Linholm, "Role of Test Chips in Coordinating Logic and Circuit Design and Layout Aids for VLSI," *Solid State Technology*, vol. 24 (9), pp. 68-74, Sept. 1981.
3. M. G. Buehler, "Comprehensive Test Patterns with Modular Test Structures: The 2 by N Probe-Pad Array Approach", *Solid State Technology*, vol. 22 (10), pp. 89-94, Oct. 1970.
4. M. G. Buehler, T. W. Griswold, C. A. Pina, B. R. Blaes, C. C. Timoc, R. H. Nixon, S. F. Suszko, "Product Assurance Technology for Procuring Reliable, Custom LSI/VLSI Electronics", *JPL Publication*, p. 83-70, Sept. 1983.
5. COS/MOS Integrated Circuits Manual, RCA Technical Series CMS-271 (1972).
6. C. F. Hill, "Noise Margin and Noise Immunity in Logic Circuits", *Microelectron.*, vol. 1, pp. 16-21, Apr. 1968.
7. M. G. Buehler, T. W. Griswold, "The Statistical Characterization of CMOS Inverters Using Noise Margins," *Electrochem. Soc. Ext. Abst.*, No. 257, pp. 391-392, May 1983.
8. S. M. Kang, "A Design of CMOS Poly Cells for LSI Circuits", *IEEE Trans. on Circuits and Systems*, vol. CAS-28, pp. 838-843 (1981).

Martin G. Buehler received his B.S.E.E. and M.S. degrees from Duke University and in 1966 received a Ph.D. degree in Electrical Engineering from Stanford University. Prior to joining the National Bureau of Standards in 1972, he was a staff member of the Semiconductor Research and Development Laboratories, Texas Instruments, where he applied electrical measurement techniques to the detection of defects and the profiling of dopants in semiconductor materials. At NBS he led a group in process metrology, designed nearly a dozen semiconductor test structures, was awarded the Department of Commerce's Silver Medal for creative contributions to semiconductor metrology, and initiated a program on self-test techniques for VLSI. Dr. Buehler joined the Jet Propulsion Laboratory in 1981 and is currently the principal investigator of the Product Assurance Technology Program for LSI/VLSI and the GaAs IC Design and Product Assurance Programs. He is a member of Tau Beta Pi, Eta Kappa Nu, Sigma Xi, the Electromechanical Society, and IEEE.

Hoshyar Sayah received a B.S. degree in Electrical Engineering from the University of California at Los Angeles in 1983. Since 1983, he has been with Caltech Jet Propulsion Laboratory, working on the development of test structures for characterizing integrated circuit process faults.

2.3.4 Reliability Analysis

N86-29260

The objective of this effort is to search for and demonstrate approaches and concepts for fast wafer probe tests of mechanisms affecting the reliability of MOS technology and, based on these, develop and optimize test chips and test procedures.

Here we report our progress on four important wafer-level reliability problems:

(1) Gate-oxide radiation hardness.

(2) Hot-electron effects.

(3) Time-dependent dielectric breakdown.

(4) Electromigration.

2.3.4.1 Gate-Oxide Radiation Hardness. The conventional method of determining the total-dose radiation hardness of MOS gate oxides involves exposing the device to ionizing radiation under an appropriate bias and measuring of the consequent threshold voltage and transconductance changes. Since the radiation response of the gate oxide is strongly process-dependent, it is desirable to routinely monitor the gate hardness in a process line. Such evaluations using conventional radiation sources such as Co^{60} cannot be carried out physically close to the process line and are cumbersome, time-consuming, and expensive. In the following, we will describe a technique we have developed to simulate the radiation environment by purely electrical means along the lines proposed by Boesch and McGarrity [1]. With this technique it is possible to make fast measurements of gate oxide radiation hardness at the wafer probe level, along with any set of other electrical measurements.

The physical principles involved are depicted in Figure 2.3.4-1. This figure shows the energy-band diagram of an n-channel transistor under a high electric field (top) and high energy radiation (bottom). Under the high field the electrons are injected into the oxide via the Fowler-Nordheim (FN) tunneling mechanism. A small fraction of electrons gain enough energy from the field to generate electron-hole pairs by impact ionization. The holes drift back towards the $Si-SiO_2$ interface. Some of the holes get trapped in the bulk and the interfacial region, some interact with the interface and generate interface traps, and the rest drift into the silicon. The same description applies in the case of ionizing radiation except that in this case the holes are generated by the high energy particles. The close similarity between the two cases suggests that these stresses produce similar damage in the oxide, as will be shown to be the case. There are also differences, namely, the hole/electron population ratio in SiO_2 is unity for radiation compared to 10^{-3} - 10^{-2} for FN injection, and also the fields are much different. The large number of electrons in the FN injection case leads to larger recombination of electrons with the trapped holes. Therefore, in the FN case, a somewhat smaller change in the threshold voltage should result for the same level of hole injection.

Figure 2.3.4-2 shows the capacitance-voltage (C-V) curves for the gate oxide of a p-channel transistor. The heavy curve is the initial dependence, while the other curves were measured after injection of holes to the levels indicated. The response is recognized to be similar to the radiation effects. To make a quantitative comparison, threshold voltage shifts have been plotted in Figure 2.3.4-3 as a function of the hole fluence for radiation and Fowler-Nordheim injection. The results are in agreement; the FN shifts are somewhat smaller, as predicted by theory. A brief description of the experimental technique is given below. A detailed treatment of the subject is given elsewhere [2].

The experiments were performed on test transistors included on test chips that were processed at a commercial 3-μm silicon-gate CMOS-LSI fabrication line. The process was not radiation-hardened and thus was expected to be sensitive to ionizing radiation (or, equivalently, to hole injection stress). The gate oxidation step was carried out in normally dry oxygen with 3% HCl at 1000°C, and the gate oxide was grown to a thickness of about 85 nm. The test transistors included both p- and n-channel types of various sizes. The largest transistors had circular ring-structure gates (the outside diameter was 220 μm and the inside diameter was 180 μm) and were used to increase the sensitivity of the current/charge measurements. The smaller transistors (inside the array) had more conventional rectangular structures ranging down to a minimum size of 5-μm channel length. In addition to the obvious substrate doping difference between the n- and p-channel transistors, the gate oxide of the p-channel transistor was exposed to an ion implant (for the purpose of threshold voltage adjustment). Otherwise, the n- and p-channel transistors were processed using identical techniques. We obtained an accurate measure of the impact ionization coefficient. Using these data, we found an excellent agreement with the effect of the hole fluence equivalently released into the oxide by ionizing radiation.

Figure 2.3.4-1 depicts the energy band picture of the Fowler-Nordheim tunnel injection and impact ionization used in this work. A large positive bias ($\gtrsim$70 V or > 8 MV/cm) is applied to the poly Si gate of either n- or p-channel devices for a few seconds. During this time, an appreciable electron current is injected. These electrons are accelerated by the applied field and some fraction of them reach an energy that causes impact ionization, generating electron-hole pairs in the oxide. The holes thus generated will drift with the field back towards the silicon interface (cathode) with some fraction being trapped in the oxide. The positive space charge produced by these trapped holes enhances the field at the cathode and increases the electron injection current. If this positive feedback process is allowed to continue, breakdown ultimately occurs. By limiting the fluence of injected electrons (and thus holes generated), we obtain the information we seek without encountering breakdown.

The hole flux returning to the silicon interface can be directly measured by the carrier separation technique of Weinberg [3], as indicated by Figure 2.3.4-4. This measurement must be done on n-channel transistors so that the hole current can be extracted from the p-well. In practice, we measure the integrated electron and hole currents (or charge) on Keithley electrometers set in the coulombmeter mode. At the lower applied fields (~7.5 MV/cm), where the hole fluence generated is very small, it is necessary to correct for the small background current of the Keithley and the device p-n junction. In general, the hole fluence measured by this technique neglects the holes that are trapped and recombine with electrons in the oxide (the holes that are trapped and survive

recombination can be included as a correction from the measured displacement charge, $C\Delta V$). Most of the trapped holes recombine in the bulk of the oxide because of the high ratio of the electron to hole flux and the high recombination cross section 9. In the tunnel barrier region of the oxide (35 A) there is no electron flux in the conduction band to contribute to electron-hole recombination; however, a significant fraction of these trapped holes can also recombine with electrons that tunnel from the silicon interface. The relative rates of trapping and tunnel recombination will determine the final density distribution of the surviving trapped holes in the barrier region. Therefore, the total trapping-recombination loss of holes can be appreciable, and for the "rad soft" oxides is typically around 50%. However, this loss is approximately offset by a factor of two gain because the holes reaching the silicon are sufficiently energetic (their energy is at least the hole barrier height of 5 eV) to create one electron-hole pair in the silicon by impact ionization [10]. These two offsetting effects leave an uncertainty in our measure of the actual hole fluence generated in the oxide of about 20%, which is quite adequate for our purposes.

The entire sequence of electrical stress and measurements is carried out automatically using an LSI-11 minicomputer. Standard types of instruments were used, including a digital voltage supply, integrating digital voltmeter, Boonton capacitance meter and Keithley electrometers. For the low values of charge measurements, lead isolation must be on the order of 10^{14} ohms or better. The impact ionization coefficient α versus the oxide field was first determined using the n-channel transistors. The results are given in Figure 2.3.4-5. Thereafter, we performed electron tunnel injection experiments on both n- and p-channel devices by measuring the electron fluence and calculating the corresponding hole fluence generated from the predetermined value of α. Electron tunnel injection was carried out with applied fields of about 8.5 MV/cm, which gave injected current densities of the order of 10^{-6} A/cm^2. The injection was programmed to continue until a preselected electron fluence was reached (typically around 10^{15} cm^{-2}), which typically required a few seconds. The C-V curve was measured before and immediately after each injection stress. The C-V curve was then measured repeatedly at later times and under different bias conditions following injection stress, to study the relaxation behavior of charge and interface trap build-up.

2.3.4.2 Hot-Electron Effects. In recent years, with the push to attain higher and higher densities on VLSI chips, the dimensions of individual transistors have been shrinking, resulting in micron and sub-micron transistor channel lengths. At the same time, for the sake of compatibility with the existing devices, the operating voltages have generally been maintained at 5 volts. Under these conditions a very high field is produced near the drain of the transistor. In traversing this region the channel electrons heat up to an extent that they can interact with the $Si-SiO_2$ interfaces, resulting in the degradation of transistor characteristics. At present, this phenomenon is one of the major limitations in VLSI. We have reviewed the literature [4, 5, 6] extensively. There has been a great deal of interest in this problem recently, and extensive research efforts are under way at this time. A clear understanding of the processes leading to the hot-electron degradation has not yet been reached. However, an empirical model has been developed that provides an adequate basis for characterizing the behavior. Figure 2.3.4-6 shows an n-channel transistor biased in a conducting state. In short-channel transistors, the

electron population in the high field region near the drain heats up, with some electrons gaining enough energy to get injected into the oxide, producing the gate current I_G. Some high-energy electrons lose their energy by impact ionization, with the resulting holes being collected by the substrate. This is measured as I_{sub}. In Figure 2.3.4-7 we have plotted I_{sub} and I_G as functions of the gate voltage V_G. Also plotted is a variable that represents the degradation of the transistor characteristics (threshold voltage shift and transconductance change). It is observed that all these plots show a strong dependence on V_D. It is interesting that transistor degradation is strongly correlated with I_{sub} and not with I_G. This strong correlation between I_{sub} and transistor degradation provides a basis for the empirical model. The following relationships have been empirically established [5]

$$\Delta(\Delta V_{TH} \text{ or } \Delta Gm) = At^n \tag{1}$$

$$A = A_o \exp\ (-\alpha/V_D) \tag{2}$$

$$I_{sub}^m = I_o \exp(-\beta/V_D) \tag{3}$$

where t is the elapsed time with drain voltage at V_D. If Δ_{max} is the maximum degradation we can tolerate, then, combining the above relations, we get

$$\text{Failure time} = \left(\frac{\Delta_{max}}{A_o}\right)^{1/n} \left(I_o^{\ell}\ I_{sub}^m\right)^{-\ell}, \quad \ell = -\frac{1}{n}\left(\frac{\alpha}{\beta}\right) \tag{4}$$

To predict the failure time of a device, one has to determine the parameters A_o, I_o, n, α, and β by stressing the devices at a few values of V_D higher than the operating voltage. Higher drain voltages reduce the required measurement time. Once these parameters have been determined and we have set a maximum tolerable degradation Δ_{max}, then we can calculate the failure time of the device from the above relationship. I_{sub}^m in Eq. (4) is the substrate current measured at the operating drain voltage. Since most of the measurements irreversibly degrade the transistor, a number of identical transistors are required for this evaluation.

We made measurements on test transistors on test chips fabricated on commercial 3-μm silicon-gate CMOS-LSI wafers described previously. We also made measurements on 1.25-μm process test-chip transistors with channel lengths ranging from 1.0 to 1.5 mm.

A schematic illustration of the measurement apparatus is shown in Figure 2.3.4-8. A computer-controlled data acquisition system is used to run the experiments. To characterize the hot-carrier degradation, gate and substrate currents are measured under varied stress conditions of VD, VG, and (stress) time. The resulting degradation of the device is commonly given in terms of the threshold voltage shift, ΔV_{TH}, and the change in transconductance, ΔGm. The transconductance is measured in the triode region at VD = 100 mV. V_{TH} is obtained by taking the tangent to the ID vs VG curve (VD = 100 mV) at the point of maximum transconductance and extrapolating it back to ID = 0.

Our main objective in these preliminary measurements is the testing and verification of the empirical model presented in Section 2.3.4.1. In these experiments we have observed the power law dependence in Eq. (1) and the exponential dependence of Eqs. (2) and (3). Further experimentation is required for final verification of the model as a suitable tool for the evaluation of hot-electron degradation in VLSI MOSFETs.

Transistors on 3-μm and 1.25-μm fabrication line test-chips were used in these experiments. Transistors on the 3-μm line test-chips with minimum gate lengths on the order of 3 μm showed no degradation even after stress at drain voltages as high as 25 V. It is evident that hot-electron effects impose no limitation on the operation of these devices. On the 1.25-μm line test transistors (channel length 1.0 to 1.5 μm), we did observe degradation of device characteristics after stress. The degradation, however, was entirely in the reduction of transconductance. No shift in threshold voltage was observed in contrast to the results reported in the literature, which showed threshold shifts. This implies that the effect of the stress is to generate interface traps only, leaving the net surface charge at the interface at the onset of the inversion unaffected. Since the response of the gate oxide is very much process-dependent, this could be a characteristic of this particular process.

In summary, we have devised and carried out experiments with results in good agreement with the model. We will follow up on these experiments for further testing of the model and determination of a minimum number of transistors required for this evaluation. This will ultimately lead to a test procedure implemented on the automatic test system for the evaluation of test chips.

2.3.4.3 Time-Dependent-Dielectric Breakdown (TDDB). The question of dielectric integrity of the gate oxide in MOS transistors has been of great interest for a long time. With the advent of VLSI requiring gate-oxide thicknesses on the order of 200 nm or less with the concomitant high-oxide fields, development of a reliable and fast technique for TDDB evaluation has become even more important to MOS technology. We have extensively studied the breakdown mechanism in metal-gate MOS structures in the past [7, 8, 9], and in the course of the present effort we have looked extensively at the literature reviewing the results of the more recent works on poly-gate MOS structures [10, 11]. In most of the studies the structures used for TDDB evaluation are capacitor structures. In view of the fact that there are known failure modes in gate oxides associated with the edges parallel to the channel [12] as well as those edges perpendicular to the channel (source and drain) [13], it is very important that the test structure replicate as closely as possible the structure being evaluated. To that end we have developed test structures with the relative dimensions of a typical transistor. This is shown schematically in Figure 2.3.4-9. One of the

two other structures shown in this figure enhances the source-drain edge, and the other enhances the thin-thick oxide. These structures are designed to help separate the failure rates associated with the different components (area and the edges) of the device. Figure 2.3.4-10 is the photograph of the test structure that we have arrived at after some iterations. It can be of any of the three configurations shown in Figure 2.3.4-9.

A test station with a 40-point probe card and other necessary electronic support has been designed and built. The probe station is integrated into a computer-controlled system and we have begun the TDDB measurements.

2.3.4.4 Electromigration. The literature on the subject was reviewed [14, 15]. The work in this field has been going on for more than fifteen years. The theoretical understanding of the subject in terms of the forces and factors leading up to this failure is fairly well-established. But the statistical nature of the problem makes it not amenable to an exact analysis. However, a model for this mechanism has been developed that has gained a wide acceptance among the workers in the field. According to this model, the median time to failure (MTF) of a strip of metallization is given by

$$MTF = A\ J^{-n} \exp(\varphi/kT)$$

where J is the current density, T the absolute temperature, and k the Boltzmann's constant. Parameter A depends on the sample geometry, material characteristics of film and substrate, and protective coating, φ is an activation energy, and n is an experimentally determined factor that depends on the temperature.

This model will be used in the evaluation of our test structures. There are test vehicles on our present test chips for testing metal continuity that can be used in our initial experiments on electromigration as well as resulting in a more efficient use of the chip area. However, the development of optimized test structures is of utmost importance, and we shall devote a good deal of effort towards that goal.

2.3.4.5 References.

1. H. E. Boesch and J. M. McGarrity, "An Electrical Technique to Measure the Radiation Susceptibility of MOS Gate Insulators," IEEE Trans. Electron Devices, NS-26, 4814 (1979).

2. J. Maserjian and N. Zamani, "Study of Oxide Trapping and Interface Trap Generation in LSI Gate Oxides Using Fowler-Nordheim Tunnel Injection and Impact Ionization" (in press).

3. Z. A. Weinberg, "Hole Injection and Transport in SiO_2," Appl. Phys. Lett., 27 (8), 437 (1975).

4. T. H. Ning, P. W. Cook, R. H. Dennard, C. M. Osborn, S. E. Schuster, and H. N. Yu, "1-μm MOSFET VLSI technology: Part IV. Hot-Electron Design Constraints," IEEE Trans. Electron Devices, ED-26, 346 (1979).

5. K. K. Ng and G. W. Taylor, "Effects of Hot-Carrier Trapping in n- and p-Channel MOSFET's," IEEE Trans. Electron Devices, ED-30, 871 (1983).

6. E. Takeda and N. Suzuki, "An Empirical Model for Device Degradation Due to Hot-Carrier Injection," IEEE Electron Device Lett., EDL-4, 111 (1983).

7. S. P. Li and J. Maserjian, "Effective Density of Defects for MOS Breakdown: Dependence of Oxide Thickness," IEEE Trans. Electron Devices, ED-23, 525 (1976).

8. S. P. Li, E. T. Bates, and J. Maserjian, "Time Dependent MOS Breakdown," Solid-State Electron., 19, 235 (1976).

9. S. P. Li and J. Maserjian, "Prediction of Time-Field-Dependent Gate-Oxide Breakdown in MOSFET Devices," Solid-State Electron., 22, 939 (1979).

10. D. L. Crook, "Methods of Determining Reliability Screen for Time Dependent Dielectric Breakdown," Proc. Reliability Physics Symposium, 1 (1979).

11. E. Domanque, R. Rivera, and C. Shepard, "Reliability Prediction Using Large Area MOS Capacitors," Proc. Reliability Physics Symposium, 140 (1984).

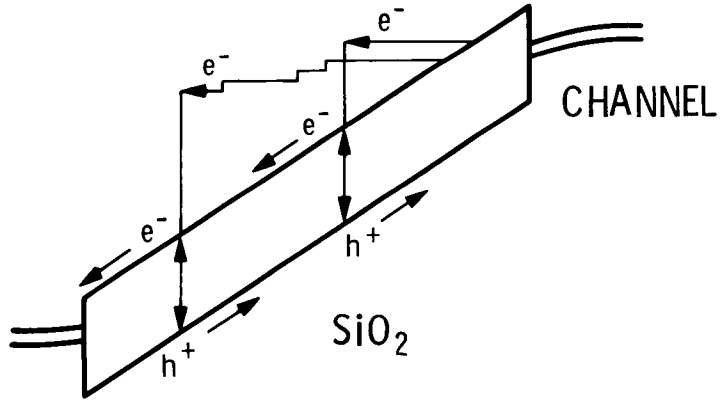
12. E. Kooi, J. G. Van Lierop, and J. A. Appels, "Formation of Silicon Nitride at a Si-SiO_2 Interface During Local Oxidation of Silicon and Heat-Treatment of Oxidized Silicon in NH_3 Gas," J. Electrochem. Soc., 123 (7), 1117 (1976).

13. S. N. Shabde, G. Simmons, A. Baluni, and D. Back, "Snapback Induced Gate Dielectric Breakdown in Graded Junction MOS Structures," Proc. Reliability Physics Symposium, 165 (1984).

14. P. B. Ghate, "Electromigration Induced Failures in VLSI Interconnects," Proc. Reliability Physics Symposium, 292 (1982).

15. T. A. Burkett and R. L. Miller, "Electromigration Evaluation--MTF Modeling and Accelerated Testing," Proc. Reliability Physics Symposium, 264 (1984).

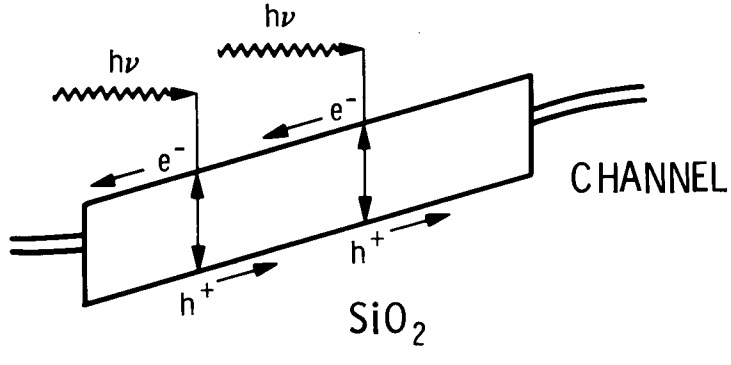
HOLE GENERATION BY F-N INJECTION/IMPACT IONIZATION



$F_{OX} \sim 8\text{-}9$ MV/cm

$\frac{N_h}{N_e} < 10^{-2}$

HOLE GENERATION BY RADIATION



$F_{OX} \sim 1\text{-}2$ MV/cm

$\frac{N_h}{N_e} = 1$

Figure 2.3.4-1. Electrical test of radiation hardness

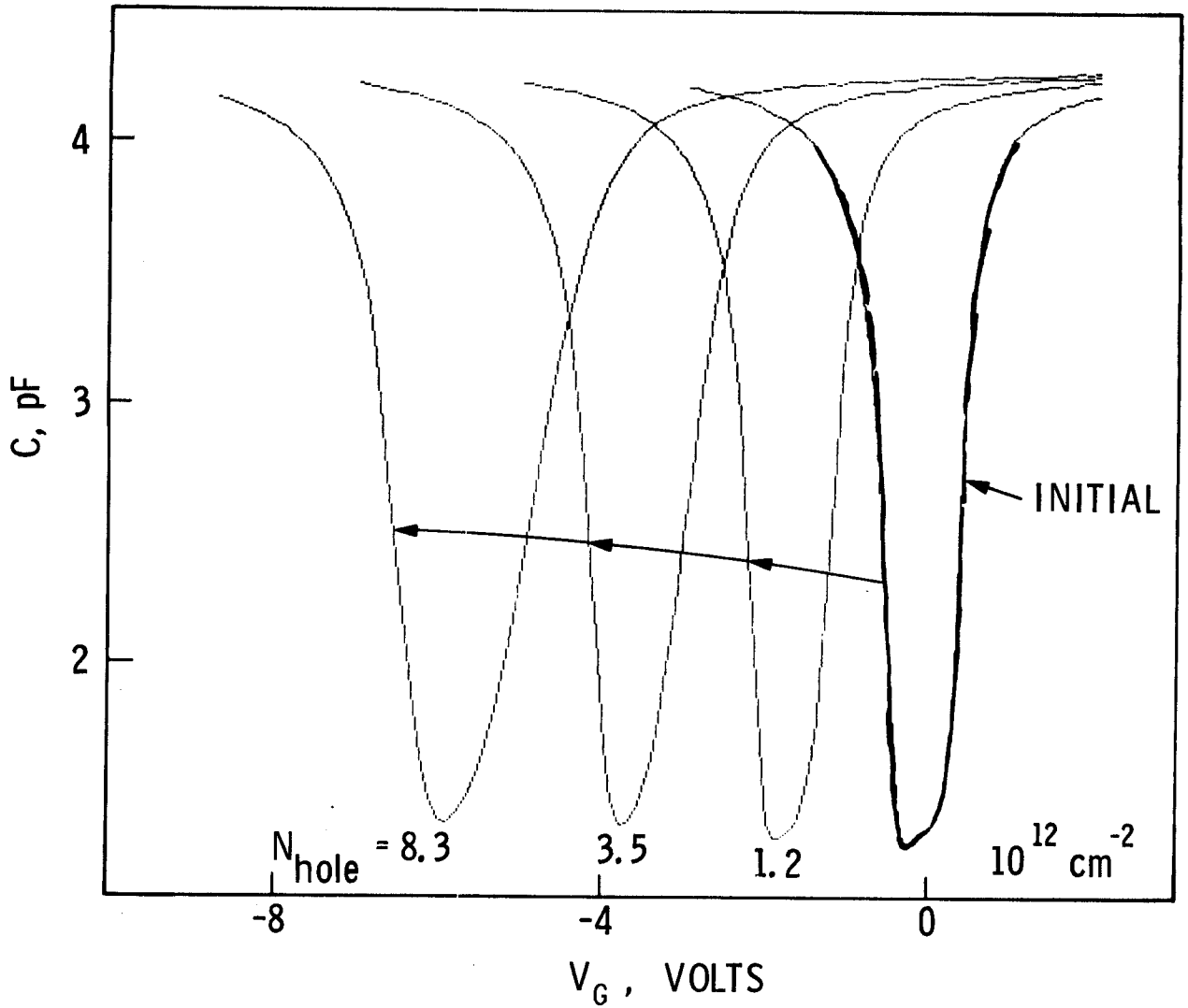


Figure 2.3.4-2. C-V shift vs injection

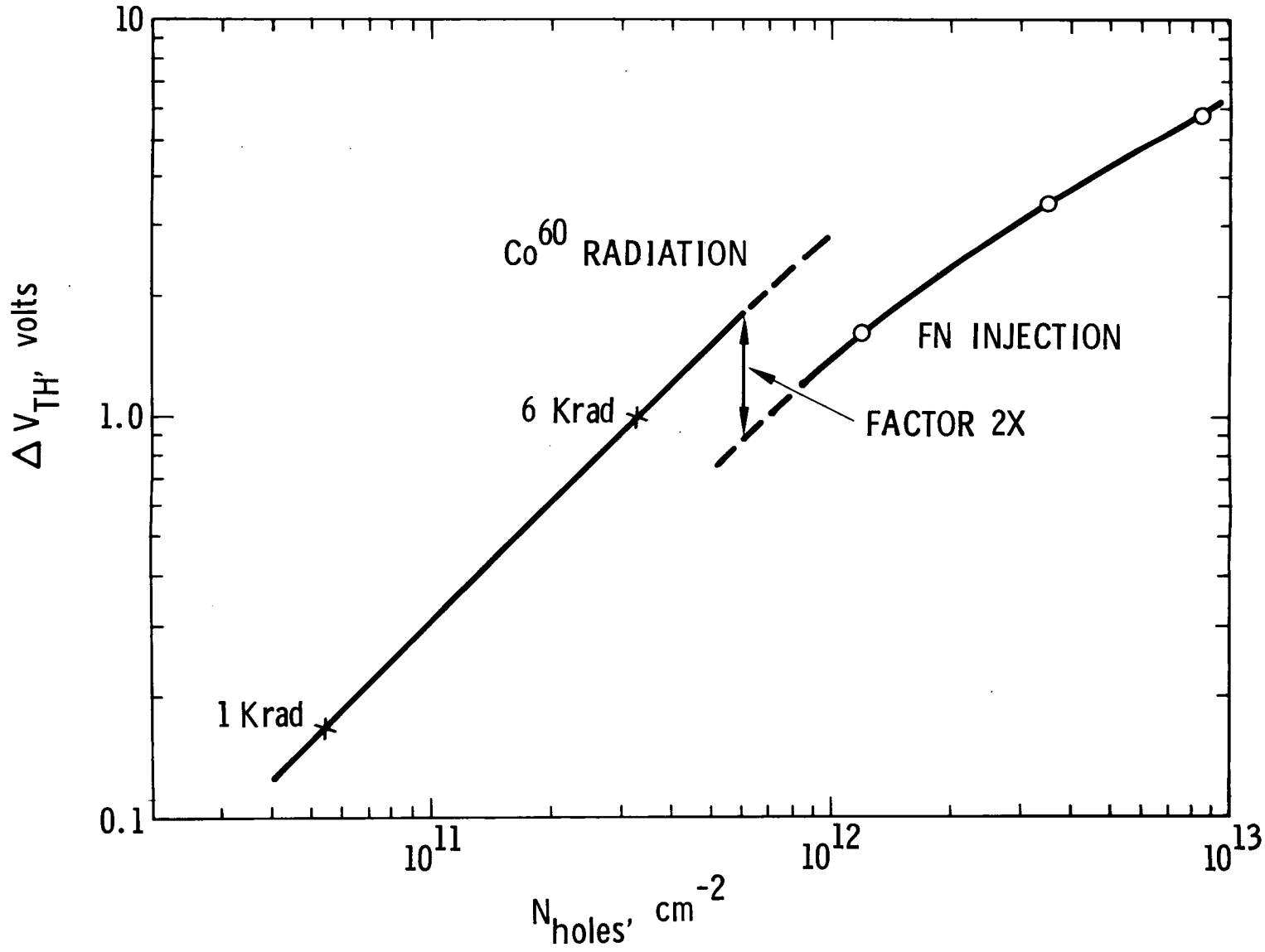


Figure 2.3.4-3. Comparison of radiation vs FN injection

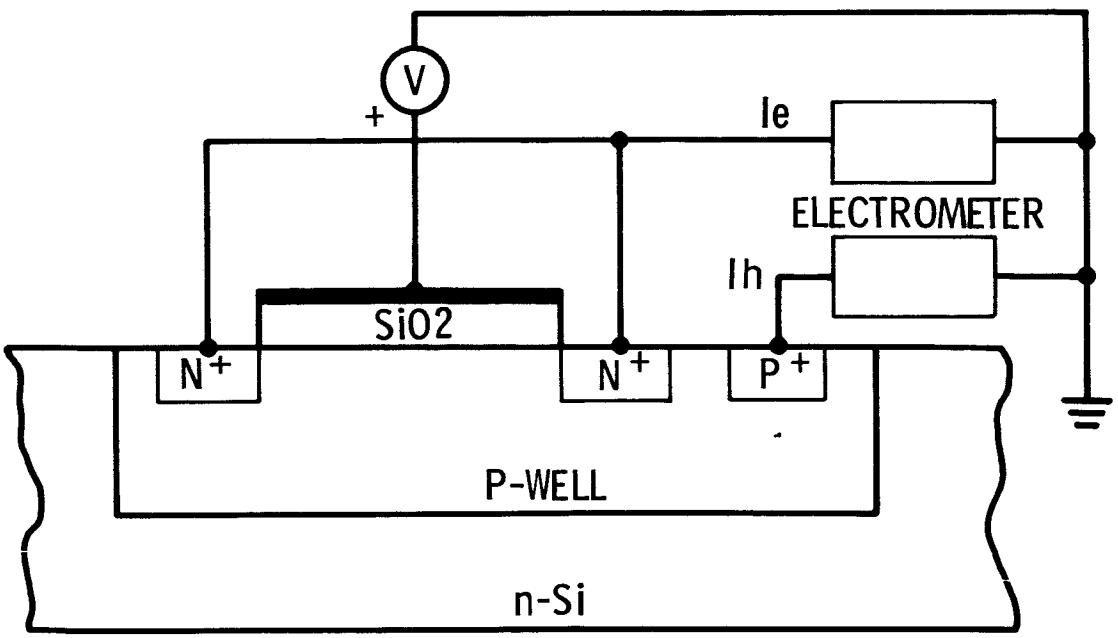


Figure 2.3.4-4. Impact ionization measurements apparatus

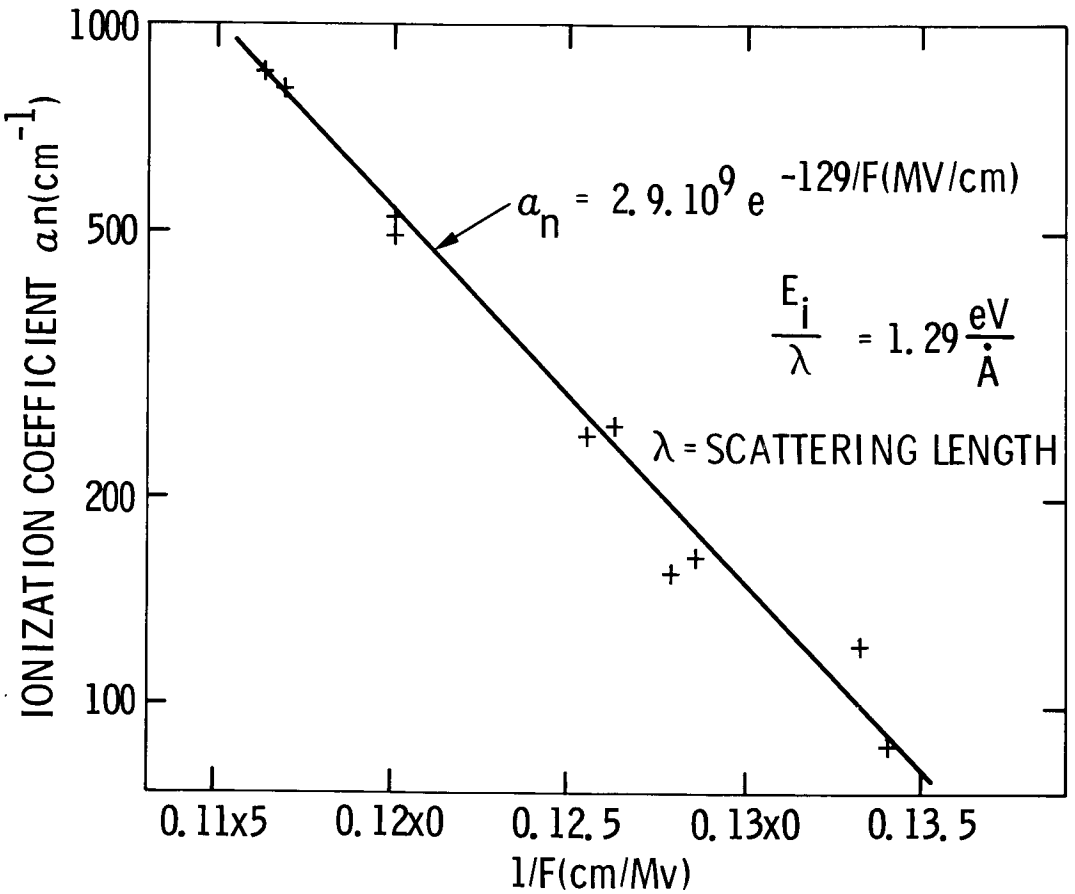
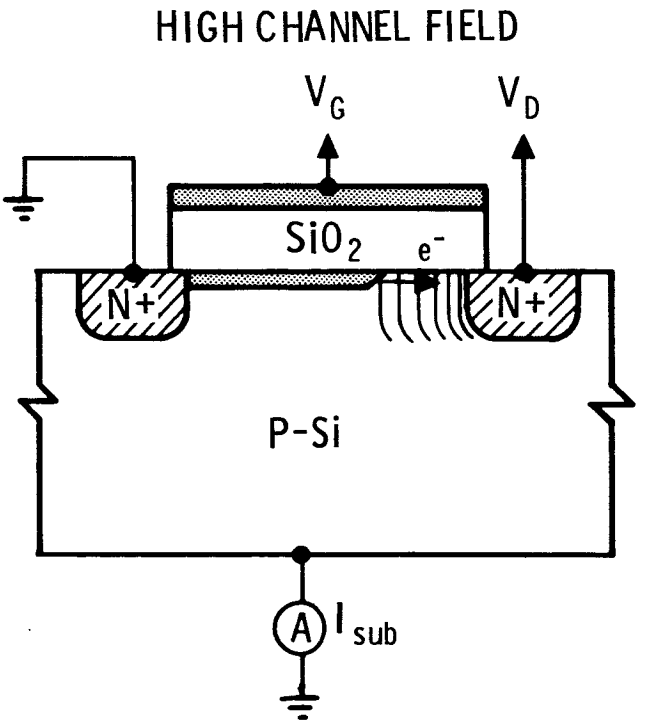


Figure 2.3.4-5. Impact ionization coefficient for electrons



EMPIRICAL MODEL

$$\text{FAILURE TIME} = \left(\frac{\Delta_{MAX}}{a_o}\right)^{1/n} I_o^{\ell} \left(I_{sub}^{m}\right)^{-\ell}$$

Figure 2.3.4-6. Hot-electron effects for an n-channel transistor biased in a conducting state

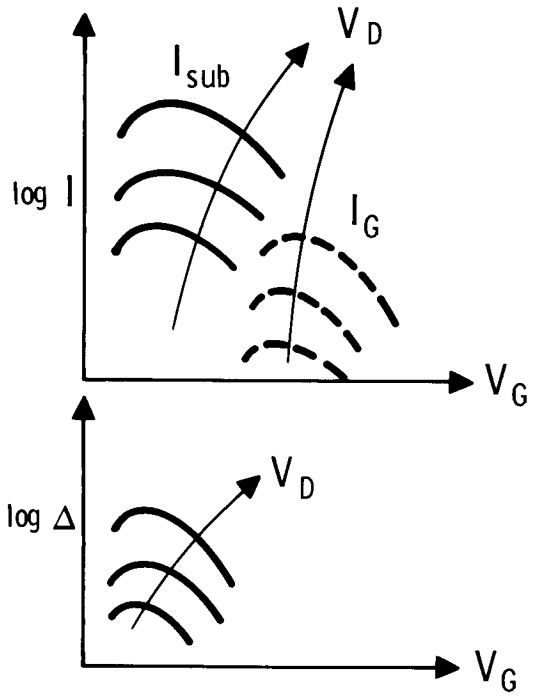


Figure 2.3.4-7. I_{sub} and I_G as functions of the gate voltage V_G

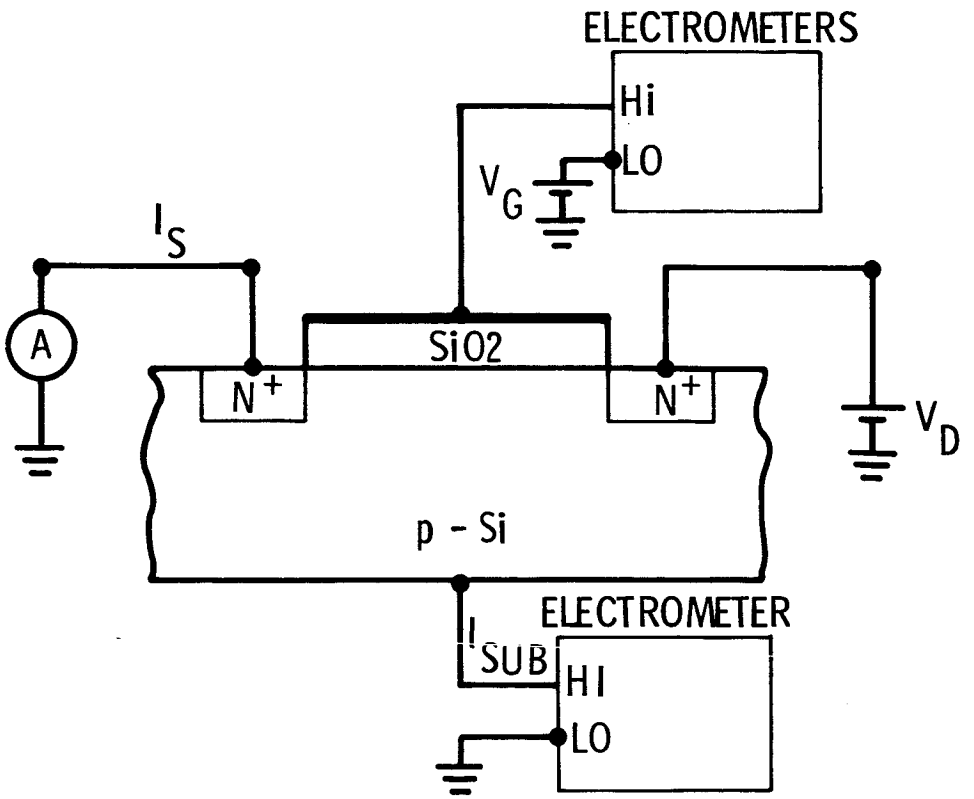


Figure 2.3.4-8. Hot-electron effects measurement apparatus

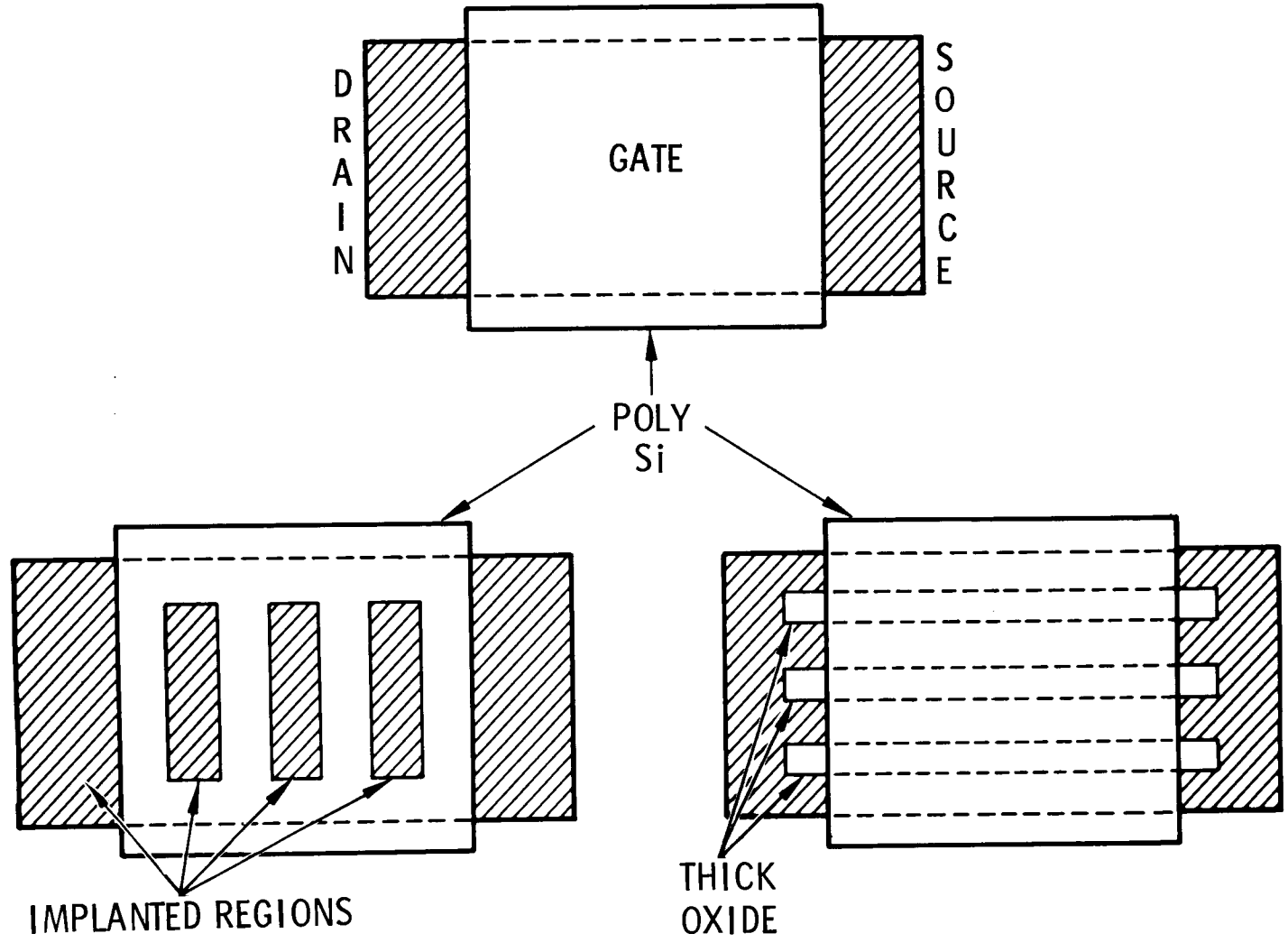


Figure 2.3.4-9. TDDB test structures

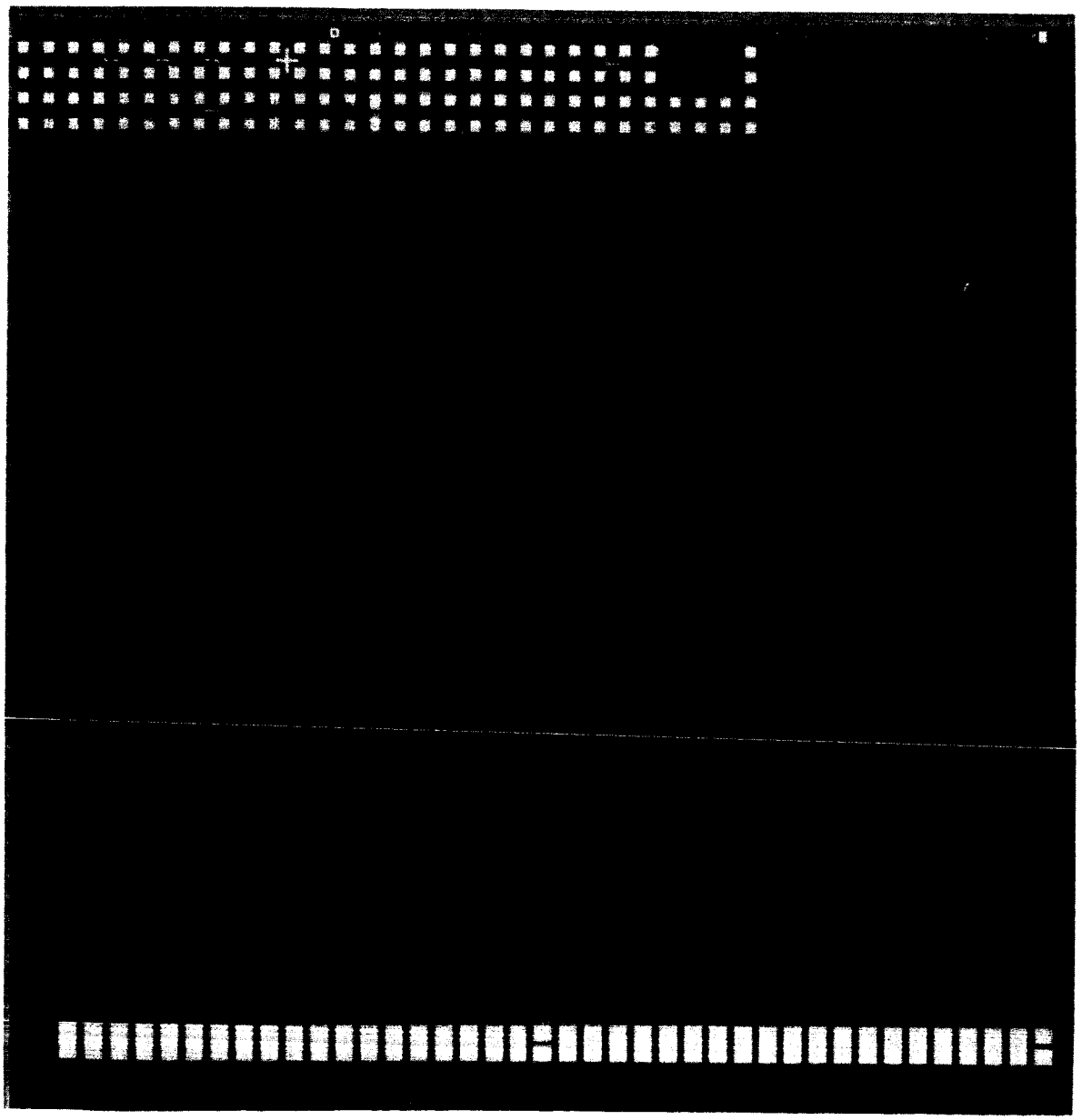


Figure 2.3.4-10. TDDB test chip

2.3.5 Inverter Noise Margin Analysis[1]

2.3.5.1 Analysis. The analysis presented early in the contract period [1] was extended to include the effect of a noise margin distribution that was not centered on the design center. In this analysis the inverter transfer curve is approximated with piecewise segments, mirrored about the Vout = Vin line and the noise margin, M, determined using the maximum square method:

$$M = (1 - 1/G)V \tag{1}$$

where G is the magnitude of the slope of the inverter transfer curve at the inverter threshold voltage, V. Gaussian statistics are used to determine the probability that one inverter has its noise margin within ± delta M of the noise margin design center, Mc:

$$P(XL < X < XU) = (1/SQR(2 \cdot Pi)) \int_{XL=X2-X1}^{XU=X2+X1} \exp((-X^2)/2)dX \tag{2}$$

where

$X = (M - Mu)/Ms$,

$X1 = \text{delta } M/Ms$,

$X2 = (Mc - Mu)/Ms$,

$Mu = (1 - 1/Gu)Vu = Vu - X3$, and

$X3 = Vu/Gu$.

In the above relationships, Mu is the mean noise margin, Vu is the mean inverter threshold voltage, Ms is the noise margin standard deviation, and Gu is the mean inverter gain.

The probability that N inverters have their noise margins within ± delta M of Mc is:

$$P(N) = P^N \tag{3}$$

where P is given in Eq. (2) above. This relationship follows from the assumption that each inverter is statistically independent of the others [2].

In the previous analysis [1], the probability that N inverters are within ± delta M was shown in Figure 4 of [1] for X2 = X3 = 0. In this

[1]This section was prepared by M. G. Buehler.

analysis,Figure 2.3.5-1 plots the number of inverters per circuit that have a probability of 50 percent of having all their inverters within ± delta M for X3 = 0. Note that the X2 = 0 curve can be derived from Figure 4 shown in the previous analysis for P = 0.5. Figure 2.3.5-1 shows that the number of inverters per circuit can range from 1 to 1,000,000, depending on whether the noise margins are centered (X2 = 0) or are tightly distributed (Ms = small). If both of these conditions are met, then one can expect high yielding circuits.

A further analysis of the noise margin equation, Eq. (1), is indicated in Figure 2.3.5-2. Here it is seen that the noise margin equation, Eq. (1), is restricted to Vl < V < VDD/2. For V > VDD/2 the noise margin equation is:

$$M = (1 - 1/G) \bullet (VDD - V) \qquad (4)$$

where VDD is the voltage supplied to the inverter. The situation is further clarified in Figure 2.3.5-3, where Eqs. (1) and (4) are plotted for various values of G. This figure further illustrates that for G > 10, the error in calculating the noise margin, assuming G is infinite, is less than 10 percent.

2.3.5.2 References.

1. M. G. Buehler and T. W. Griswold, "The Statistical Characterization of CMOS Inverters Using Noise Margins," Electrochemical Society Extended Abstract 257, 391-392 (May 1983).

   Note: This extended abstract is reprinted in its entirety as it appeared in the journal, following Figure 2.5.3-3.

2. S. L. Meyer, Data Analysis for Scientists and Engineers, J. Wiley and Sons, New York (1975).

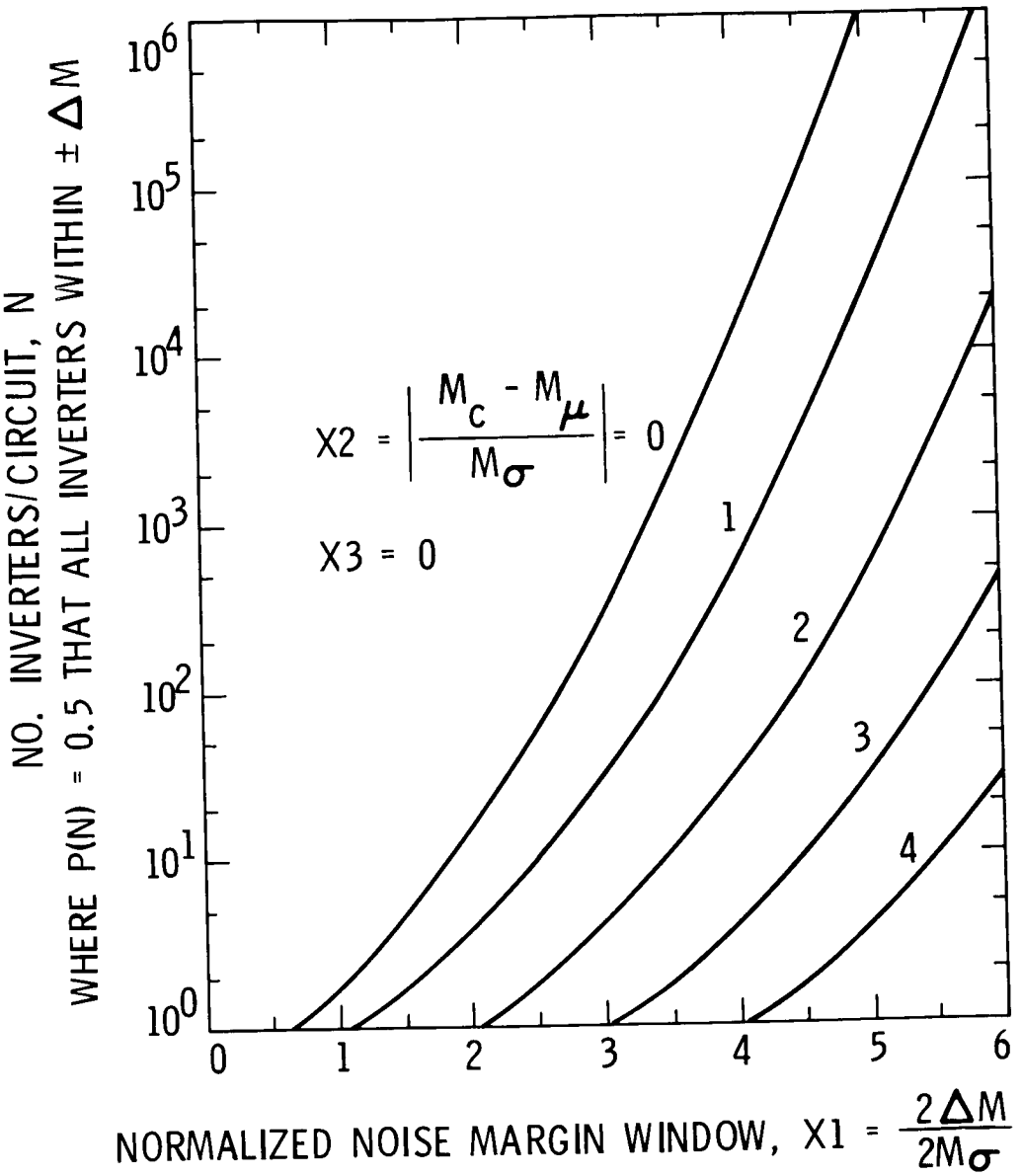


Figure 2.3.5-1. The number of inverters per circuit where the probability is 50 percent that all the inverters are with the design window, 2ΔM, and centered within X2 of the design center, Mc, for G = infinity

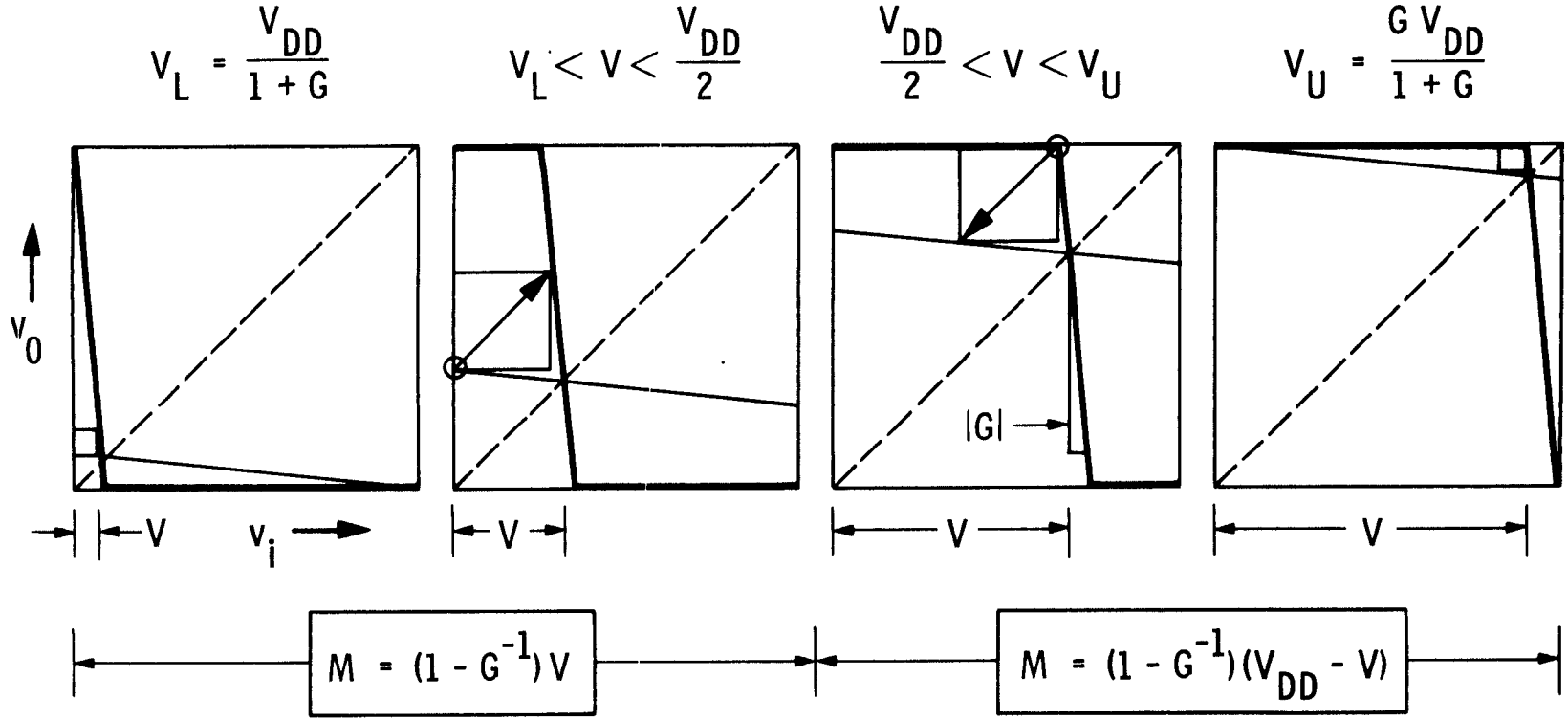


Figure 2.3.5-2. Inverter noise margin as derived from the maximum square method and illustrated for four cases

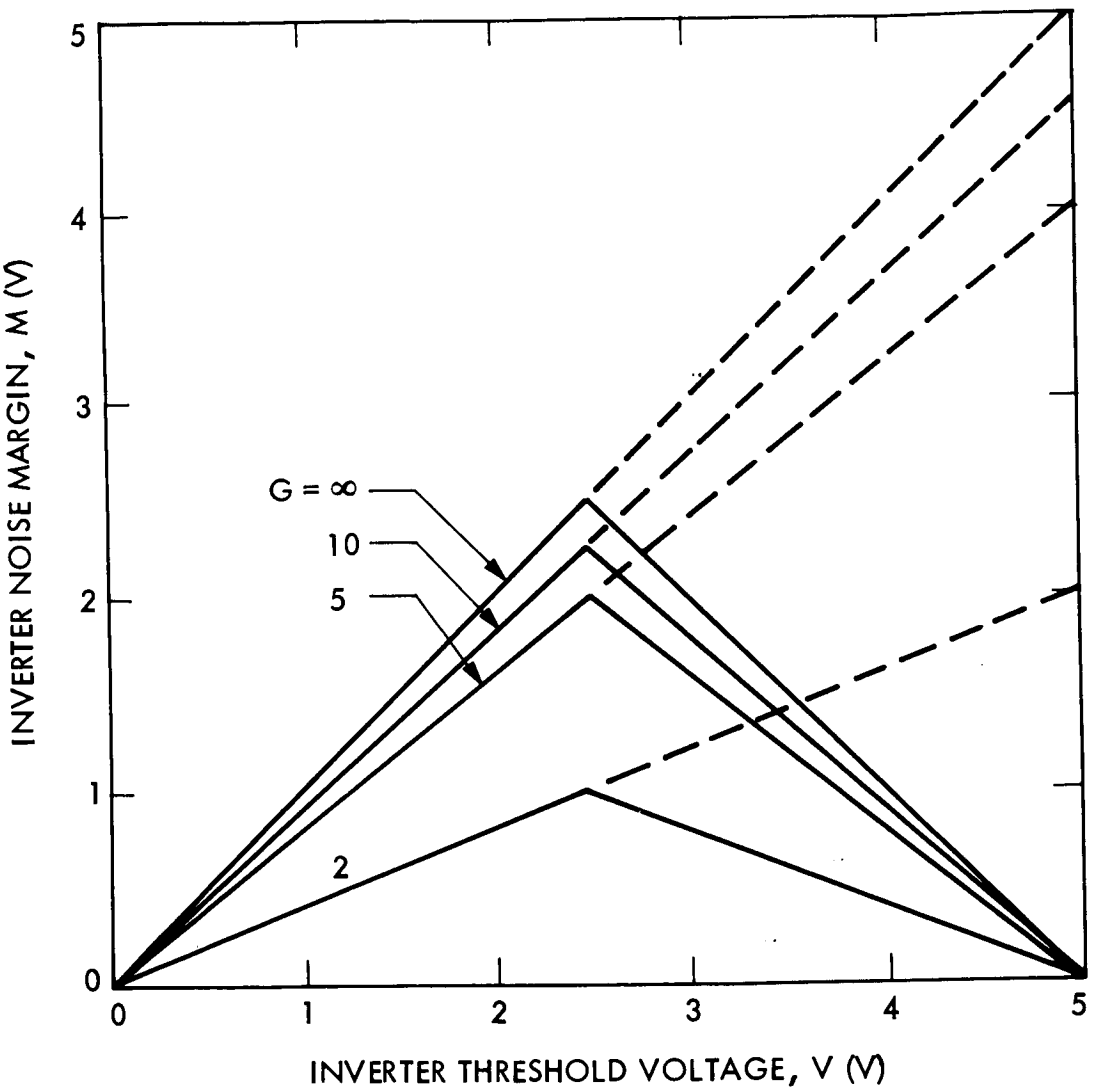


Figure 2.3.5-3. Inverter noise margin dependence on the inverter threshold voltage and gain

ABSTRACT NUMBER 257 Electrochemical Society (May 9, 1983)

THE STATISTICAL CHARACTERIZATION OF CMOS INVERTERS USING NOISE MARGINS

M. G. Buehler and T. W. Griswold
California Institute of Technology
Jet Propulsion Laboratory
Pasadena, CA 91109

Introduction

The goal of this effort is to develop simple techniques that can be used to assess CMOS bulk technology for use in critical circuit applications. In this effort test chips were developed, fabricated at silicon foundries, and evaluated using parametric test equipment. The approach is based on the statistical analysis of inverter noise margin and required the measurement and analysis of a large amount of data; this necessitated the development of computer-aided test equipment and the use of computer-aided analysis techniques. The CMOS inverter was found to be a suitable test structure for evaluating a portion of the capability of a wafer fabrication process. A procedure was developed for determining the number of inverters that can be fabricated in a circuit for a probability of 0.5 that all inverters are within a specified noise-margin window. Results from the best wafers indicate that it is possible to obtain circuits having more than 10^6 inverters all within a noise-margin window as small as 1.0V.

Experimental Approach

A p-well CMOS bulk microelectronic test strip consisting of transistors, inverters, cross-bridge resistors, and contact resistors was fabricated along with more elaborate test chips at several 5-μm silicon foundries. Data was acquired from the silicon wafers using a parametric test system. In order to assure the integrity of the wafer probe data, various procedures were employed to eliminate invalid data from the data set. Then outliers were identified using a modified version of the NBS STAT2 program [1].

The inverters used in this study have an as-drawn channel length by width of 5μm x 5μm for the n-channel transistor and 5μm x 7.5μm for the p-channel transistor. The transfer curves for 28 CMOS inverters are shown in Fig. 1 where the input voltage is v_i and the output voltage is v_o. Normally this curve is characterized by four parameters: $v_o(v_i=0)$, $v_o(v_i=V_{DD})$, V (inverter threshold voltage), and G (inverter gain) determined from the slope of the transfer curve at v_i= V±25mV.

Inverter Analysis

The CMOS transfer curve was approximated with a three-section piece-wise linear curve as shown in Fig. 2. The middle section of the curve was fitted to the slope, G, of the measured transfer curve at the inverter threshold voltage, V. The noise margin was determined from the maximum square approach of Hill [2]. In this approach the piece-wise linear curve was mirrored about the v_i=v_o curve, and the noise margin, M, derived from the diagonal line of the maximum square as illustrated in Fig. 2. The expression for the "low" noise margin is

$$M = (1-G^{-1})V. \qquad (1)$$

Statistical variations in the noise margin are illustrated in Fig. 3 where the distribution is assumed to be normally distributed with a mean, M_μ, and standard deviation, M_σ. In order to evaluate whether the M_μ and M_σ are satisfactory, they must be related to a noise margin target, M*, and window, 2ΔM. The probability of finding individual inverters within the target window is

$$P(XL<X<XU) = (2\Pi)^{-1/2} \int_{XL=X2-X1}^{XU=X2+X1} EXP(-X^2/2)dX \qquad (2)$$

where

$$X = (M-M_\mu)/M_\sigma \qquad (2a)$$

$$X1 = \Delta M/M_\sigma \qquad (2b)$$

$$X2 = (M^*-M_\mu)/M_\sigma. \qquad (2c)$$

By combining N inverters into circuits where each inverter is statistically independent, the probability of N inverters per circuit being within ±ΔM is:

$$P(N) = P^{**}N \qquad (3)$$

where P is given by Eq (2). This probability is illustrated in Fig. 4 where the noise margin target was set equal to the mean (i.e., $M^*=M_\mu$ or X2=0). These curves show that as the number of inverters increases from N=1 to 10^6, the standard deviation, M_σ, must get smaller for a fixed ΔM in order to have a good probability of finding all inverters per circuit within the noise-margin window. If the noise margin mean, M_μ, is not equal to the target, M*, then the curves shift to the right along the X1-axis by one unit for each unit change in X2.

The mean and standard deviation for the noise margin were derived from inverter gain and threshold voltage parameters. Since G and V combine non-linearly to form the noise margin, a Taylor series expansion [3] of Eq (1) was necessary to evaluate the noise margin mean:

$$M_\mu = (1-G_\mu^{-1})V_\mu \qquad (4)$$

and the noise margin standard deviation

$$M_\sigma = \left[(V_\mu^2 G_\mu^{-4})G_\sigma^2+(1-G_\mu^{-1})V_\sigma^2\right]^{1/2} \qquad (5)$$

where the mean and standard deviation for the gain are G_μ and G_σ, respectively, and for the threshold voltage are V_μ and V_σ, respectively.

Results from seven wafers are listed in Table 1. The M_μ and M_σ values were derived from the G and V parameters where invalid and outlier values were excluded from the data sets. The number of values used to compute the G and V parameters ranged from 34 in the poorest to 135 in the best wafer.

Table 2 lists the number of inverters that can be fabricated per circuit with a probability of 0.5 that all inverters are within selected noise margin windows. The target M*=2.25V was determined from ideal G and V values of 10 and 2.5V, respectively. The numbers in Table 2 were calculated from Eqs (2) and (3) using the parameters from Table 1.

From Table 2, one can quickly evaluate not only the capability of the technology but also the controllability of the wafer fabrication process. Results from the best wafers indicate that it is possible to obtain circuits having more than 10^6 inverters all within a noise-margin window (2ΔM) as small as 1.0V. Since this evaluation was based on whole wafer statistics, it is thought that variations within a chip will be less, thus enhancing the ability to fabricate large circuits.

This paper was originally presented at the Spring 1983 Meeting of The Electrochemical Society, Inc. held in San Francisco, California. Reprinted by permission from The Electrochemical Society.

TABLE 1. CMOS INVERTER PARAMETERS

WAFER	G_μ	G_σ	V_μ(V)	V_σ(V)	M_μ(V)	M_σ(V)
1202	17.6	3.33	2.31	0.206	2.18	0.196
1205	14.9	1.86	2.52	0.033	2.35	0.037
1213	17.4	1.83	2.46	0.026	2.32	0.029
2205	11.8	1.32	2.12	0.227	1.94	0.209
2207	10.7	1.43	1.97	0.198	1.79	0.181
2111	17.7	1.08	2.30	0.072	2.17	0.068
2120	10.3	1.50	2.23	0.055	2.01	0.059

TABLE 2. N INVERTERS/CIRCUIT FOR P(N)=0.5 THAT ALL LIE WITHIN ±ΔM OF M*=2.25V

	ΔM(V)		
WAFER	0.2	0.5	1.0
1202	1.6E0	4.2E1	6.4E5
1205	1.7E2	>1.0E7	>1.0E7
1213	3.0E5	>1.0E7	>1.0E7
2205	0.6E0	3.5E0	1.5E3
2207	0.3E0	1.3E0	4.4E2
2111	1.7E1	>1.0E7	>1.0E7
2120	0.5E0	1.8E5	>1.0E7

Acknowledgments

The authors wish to thank B. Blaes of JPL for acquiring the data, C. Pina of JPL for designing the test strip, and R. Zuleeg of McDonnell Douglas for pointing out the analysis of Hill [2]. This work was supported under a joint NASA/DARPA/NSA contract No. NAS7-100, Task Order #RD168, Amendment #116.

References

1. R. L. Mattis, L. J. Till, and R. C. Frisch, "A Computer Program for Analysis of Data from Microelectronic Test Structures," NBSIR 82-2492 (June 1982).

2. C. F. Hill, "Noise Margin and Noise Immunity in Logic Circuits," Microelectronics 1, 16-21 (April 1968).

3. A. H. Bowker and G. J. Lieberman, "Engineering Statistics," Prentice-Hall (New York 1959), p. 62.

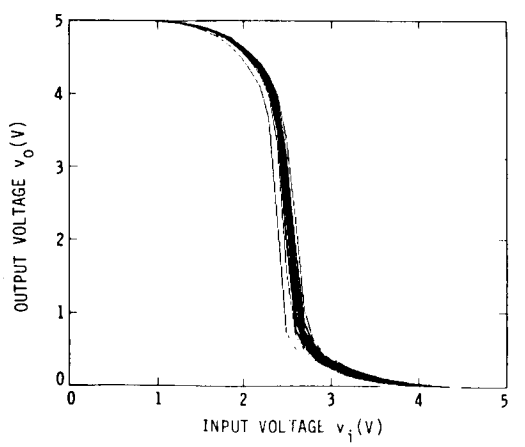


Fig. 1. Transfer curve for 28 CMOS bulk inverters measured on wafer 1205.

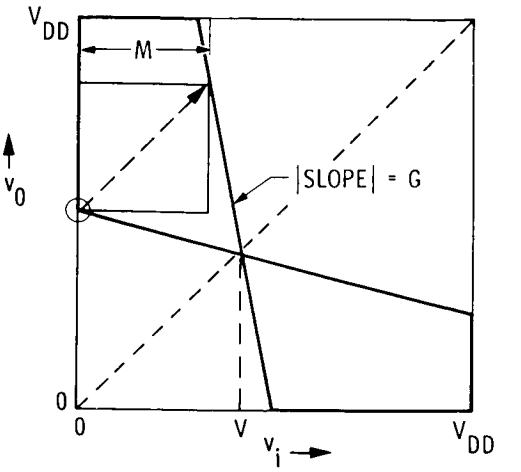


Fig. 2. Maximum-square approach to defining the noise margin, M, using a piece-wise linear fit to the inverter transfer curve.

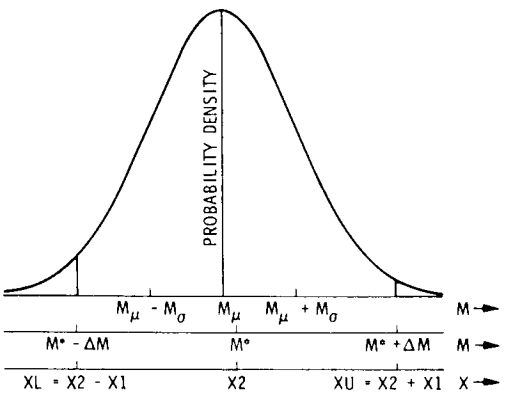


Fig. 3. Normal probability density function related to the noise margin target, M*, and noise margin window, 2ΔM.

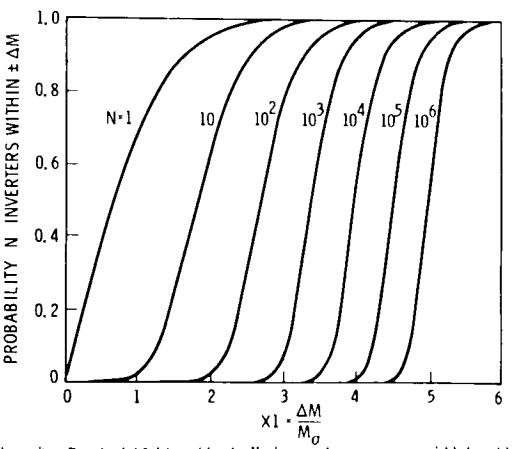


Fig. 4. Probability that N inverters are within the noise margin window, 2ΔM, as ΔM varies relative to the noise margin standard deviation, M_σ. These curves assume that $M^* = M_\mu$ so that X2=0.

2.3.6 JMOSFIT--A MOSFET Parameter Extractor with Geometry-Dependent Terms[1]

2.3.6.1 Overview

2.3.6.1.1 Introduction. This effort was motivated by the need to extract the parameters from metal-oxide-silicon-field-effect transistors (MOSFETs) that are included on the CRRES test chip (see Section 2.6), to have a simple but comprehensive method that can be used in wafer acceptance, and to have a method that is sufficiently accurate that it can be used in IC design.

The goals of this effort are to develop a set of MOSFET parameter extraction procedures that are directly linked to the MOSFET model equations and that facilitate the use of simple, direct curve-fitting techniques. In addition, we wish to include the major physical effects that affect MOSFET operation in the linear and saturation regions of operation for devices fabricated in 1.2- to 3-mm CMOS technology. The fitting procedures were designed to establish single values for such parameters as threshold voltage and transconductance and to provide for slope matching between the linear and saturation regions of the MOSFET output current-voltage curves. In this presentation we have preserved as much of the traditional SPICE [1] MOSFET terminology as possible.

Numerous approaches can be found in the literature for determining the parameters needed by the MOSFET equations [2, 3, 4, 5, 6, 7, 8]. These approaches range from the optimization techniques typified by SUXES [9] to the parameter plotting techniques described by Hanafi et al. [3]. The approach suggested here uses multiparameter linear and nonlinear least-squares fitting techniques [10, 11] to fit the equations. The geometry dependence of parameters is introduced in a manner similar to the CASMOS [6] extraction approach. In this analysis four different sizes of transistors are used that cover a rectangular-shaped region of the channel length-width plane; other approaches cover an L-shaped portion of the channel length-width plane [12].

In this approach the MOSFET equations for the linear and saturation regions are derived using first principles and are arranged to simplify the fitting procedure. The following effects are considered: narrow and short channel effects, channel mobility degradation, velocity saturation, body effect, and channel length modulation. At the transition from the linear to the saturation region at a constant gate voltage, the drain current and its first derivative with respect to the drain voltage are continuous.

2.3.6.1.2 List of Frequently Used Symbols

I	= Drain current
VD, UD	= Drain-source voltage (intrinsic, extrinsic)
VG, UG	= Gate-source voltage (intrinsic, extrinsic)
V_y	= Channel-source voltage
VB	= Body-source voltage

[1]This section was prepared by M. G. Buehler and B. T. Moore.

VT	= Threshold voltage
ϕ	= PHI = Twice the body Fermi potential
ψ	= PSI = Threshold voltage less the body effect terms
VX	= VDsat = Drain-source voltage at saturation
KP	= $\mu_o C_o$ = Intrinsic channel conduction factor
β	= BETA = Channel conduction factor
δ	= DELTA = Secondary body effect factor
η	= ETA = Drain-field mobility degradation factor
ε	= EPSILON = Velocity saturation/gate-field mobility degradation factor
γ	= GAMMA = Body effect factor
λ	= LAMBDA = Channel length modulation factor
R	= Series resistance
τ	= TAU = Theta/series resistance factor
θ	= THETA = Gate-field mobility degradation factor
L, L, LE	= Channel length (as-drawn, delta, effective)
W, W, WE	= Channel width (as-drawn, delta, effective)
Q_g, Q_c, Q_b	= Charge density (gate, channel, body)
Q_{bo}, Q_{bs}, Q_{bn}	= Body charge (uncorrected, short channel, narrow channel)
X	= Depletion width
q	= Electronic charge
ε_s, ε_o	= Dielectric constant (silicon, oxide)
N	= Body dopant density
μ, μ_o	= Channel carrier mobility (total, zero-field)
v_c	= Channel carrier drift velocity
E_y, E_c	= Transverse electric field (channel, critical)
E_x	= Channel normal electric field
C_o	= Gate-oxide capacitance per unit area
K_θ	= Theta coefficient
ζ	= VD-dependent fitting factor

2.3.6.2 Equation Summary. The equations used in this analysis are given below. For the MOSFET operating in its active region where the gate-source voltage, VG, is greater than the threshold voltage, VT, the drain current, I, in the linear region (VD $\leq$ VX and VG $\geq$ VT) is

$$I = \beta(VE - aVD/2)VD/(1 + \theta VF + \zeta VD) \quad (1)$$

where VE = VG - VT, VF = VG - ψ and a = 1 + δ. In the saturation region (VD > VX and VG $\geq$ VT), the drain current is

$$I = \beta(VE - aVX/2)VX/[1 + \theta VF + \varepsilon VX - \lambda(VD - VX)] \quad (2)$$

where the crossover drain source voltage, VX, sometimes expressed as VDsat, is

$$VX = (2VE/(ac))/[1 + (1 + 2d\varepsilon VE/(abc^2))^{1/2}] \quad (3)$$

where $b = 1 + \theta VF$, $c = 1 + \theta VE/(2ab)$, and $d = 1 - \theta/(2\varepsilon)$.

The fitting parameter ζ that appears in Eq. (1) is given by:

$$\zeta = \eta(1 - 2V_\zeta + V_\zeta^2) + \varepsilon(2 - V_\zeta)V_\zeta + (\lambda - \theta/2)(1 - V_\zeta)V_\zeta \qquad (4)$$

where $V_\zeta = VD/VX$. Imbedded in the above equations are six known parameters: I_1, the drain current; VG, the gate-source voltage; VD, the drain-source voltage; VB, the body-source voltage; L, the as-drawn channel length; and W, the as-drawn channel width.

The seven unknown parameters and their dependence on VB, L, and W are given below. The conductance factor, β, is

$$\beta = KP \bullet WE/LE = KP(W - \Delta W)/(L - \Delta L) \qquad (5)$$

where $KP = \mu_o C_o$ (μ_o is the zero-field channel mobility and C_o is the gate-oxide capacitance per unit area), $LE = L - \Delta L$ is the effective channel length, and $WE = W - \Delta W$ is the effective channel width. The threshold voltage, VT, is

$$VT = \psi + \gamma(\phi - VB)^{1/2} - KLG(\phi - VB)/LE + KWG(\phi - VB)/WE \qquad (6)$$

where ψ = PSI is a potential due to the gate-silicon work function, oxide charge, and oxide-silicon interface charge, ϕ = PHI is twice the body or bulk Fermi potential, and γ = GAMMA is the primary body effect term. The secondary body effect term is DELTA:

$$\delta = DO/(\phi - VB)^{1/2} - KLD/LE + KWD/WE. \qquad (7)$$

The series resistance and gate-field mobility degradation term is TAU:

$$\tau = \theta + KLT/LE = \theta + 2\beta R. \qquad (8)$$

where θ is the gate-field mobility degradation coefficient. The series resistance term is:

$$R = KLT/(2KP \bullet WE) \qquad (9a)$$

and the series resistance coefficient is:

$$RW = KLT/(2KP). \tag{9b}$$

The drain-field mobility degradation term is ETA:

$$\eta = HO + KLH/LE. \tag{10}$$

The velocity saturation and gate-field mobility degradation term is EPSILON:

$$\varepsilon = -EO + KLE/LE. \tag{11}$$

where $EO = \theta/2$. The channel length modulation term is LAMBDA:

$$\lambda = LO + KLL/LE - KWL/WE. \tag{12}$$

In the above expressions the geometrical coefficients are denoted in general as Kij, where i denotes a length- or width-dependent coefficient and j specifies the associated parameter. In the above equations, the offset-term is denoted as DO, θ, HO, EO, and LO, which gives the value for the individual transistor parameter for large transistor geometries.

In the short and narrow channel limit, the above equations reduce to their classical forms. These expressions have been carefully formulated so that for a long channel device ($LE \to \infty$) and for $\theta = 0$ (which leads to $\varepsilon = 0$), the drain voltage at the onset of saturation is

$$VX = VE/a \tag{13}$$

and the drain current in saturation is

$$I = \beta VE^2/(2a) \tag{14}$$

which is the square law behavior expected for a long channel MOSFET. For a short channel device ($LE \to 0$, which leads to $\varepsilon = \infty$) and for $\theta = 0$, the drain voltage at the onset of saturation is

$$VX = (2VE/(a\varepsilon))^{1/2} \tag{15}$$

and the drain current in saturation is

$$I = \beta VE/\varepsilon \tag{16}$$

which is the linear law behavior expected for a short channel MOSFET [13]. Note that ε varies between 0 and ∞. For $\theta = 0$ the value of ε indicates the degree that the transistor curves are dominated by channel carrier velocity saturation.

2.3.6.3 Theory. In this section the basic relationships are derived that were used to formulate the equation set given in the previous section. The following equations were derived for an n-channel MOSFET and are based on the assumption that the channel charge follows the gradual channel approximation and that junction leakage currents are negligible.

2.3.6.3.1 Channel Charge. An expression for the channel charge density, Q_c, was derived as follows for an n-channel MOSFET. From Gauss' law, the positive charge density (coulombs/area) on the gate, Q_g, is equated to Q_c, the negative channel charge density, and the negative bulk silicon charge density, Q_b:

$$Q_g = -(Q_c + Q_b). \tag{17}$$

From the gradual channel approximation, the positive gate charge density at a distance y from the source (see Figure 2.3.6-1) due to charge on the gate is:

$$Q_g = C_o(VG - \psi - V_y) \tag{18}$$

where C_o is the gate capacitance per unit area and V_y is the channel potential at a distance y from the source. Note that both VG and V_y are referenced to the source and are positive for the n-channel MOSFET. The PSI = $\psi = \phi + V_{fb} + V_{ion} - V_{Not} + V_{Nit}$, where ϕ = PHI is twice the body Fermi potential, V_{fb} is the flat-band potential that includes the gate-silicon work function and the as-fabricated oxide and interface charge, V_{ion} is a potential due to body-enhancing ion-implanted dopants assumed to be a delta function of dopant at the oxide-silicon interface, V_{Not} is a potential due to induced positive oxide charge [14], and V_{Nit} is a potential due to induced negative interface charge [14].

The body charge in coulombs in the absence of short or narrow channel effects is

$$Q'_{bo} = -qN \bullet LE \bullet WE \bullet X \tag{19}$$

where q is the electronic charge, N is the dopant density (cm^{-3}) in the silicon, and X is the width of the depletion region beneath the channel.

From the solution to Poisson's equation for a uniformly doped one-sided n^+p step junction, the width of the depletion region is:

$$X = [(2\varepsilon_s/qN)(V_y + \phi - VB)]^{1/2} \tag{20}$$

where ε_s is the silicon dielectric constant. Note that $X(V_y = 0) = X_s$ and $X(V_y = VD) = X_d$.

Using the approximation $(1 + x)^{1/2} \simeq 1 + x/2$ with $x = V_y/(\phi - VB)$, the Q'_{bo} term is simplified for $V_y << (\phi - VB)$:

$$Q'_{bo} = -LE \cdot WE(2\varepsilon_s qN)^{1/2}[(\phi - VB)^{1/2} + V_y/(2(\phi - VB)^{1/2})]. \tag{21}$$

The amount of bulk charge lost beneath the gate due to charge-stealing by the source and drain (see Figure 2.3.6-2) gives rise to the short channel effect, in which the threshold voltage is observed to decrease as the channel length decreases. An expression for the bulk charge lost due to the short channel effect [15] is:

$$Q'_{bs} = -(qN \cdot WE/2)(Y_s X_s + Y_d X_d) \tag{22}$$

where, as seen in Figure 2.3.6-2, the charge lost to the source and drain are approximated by triangles whose base is given by Y_s at the source and Y_d at the drain. From Yau [16] and Figure 2.3.6-2, the Y_s or Y_d is

$$Y_i = X_j\left[(1 + 2X_i/X_j)^{1/2} - 1\right] \tag{23}$$

where X_j is the source or drain junction depth. For $2X_i/X_j << 1$, $Y_i = X_i$

$$Q'_{bs} \simeq -(qN \cdot WE/2)(X_s^2 + X_d^2). \tag{24}$$

This equation was evaluated by substituting Eq. (20) for X_s and X_d. Then $2V_y$ was substituted for VD, which is justified on the basis that in a subsequent step the charge will be integrated along the channel from source to drain. That is, $\int_0^{VD} 2V_y \cdot dV_y = \int_0^{VD} VD \cdot dV_y$ so that:

$$Q'_{bs} = -2\varepsilon_s WE(V_y + \phi - VB) \tag{25}$$

The amount of bulk charge added beneath the gate due to the depletion regions formed on either side of the channel gives rise to the narrow channel effect, in which the threshold voltage is observed to increase as the channel width decreases. An expression for the bulk charge added due to the narrow channel effect [13] is:

$$Q'_{bn} = -qN \cdot LE\pi X^2/2 \tag{26}$$

where the side-wall charge was assumed to be a quarter circle, as seen in Figure 2.3.6-3(a). As discussed elsewhere [17], other geometries such as triangles and square regions also have an X^2 dependence. The devices used in this study were fabricated with an isoplanar process, and a more realistic cross section is shown in Figure 2.3.6-3(b). For this geometry Silburt et al. [2] also indicate that an X^2 dependence is appropriate. After substituting for X, the Q'_{bn} is:

$$Q'_{bn} = -\pi\varepsilon_s LE(V_y + \phi - VB) \tag{27}$$

The total bulk charge density in coulombs/cm^2 is:

$$Q_b = (Q'_{bo} - Q'_{bs} + Q'_{bn})/(LE \cdot WE). \tag{28}$$

Equations (17), (18), (19), (25), (27) and (28) were combined to give the following expression for the channel charge density:

$$Q_c = -C_o(VE - aV_y) \tag{29}$$

where $VE = VG - VT$, $a = 1 + \delta$, and

$$VT = \psi + \gamma(\phi - VB)^{1/2} - \sigma(\phi - VB) + \gamma(\phi - VB) \tag{30a}$$

$$\delta = \gamma/(2(\phi - VB)^{1/2}) - \sigma + \nu \tag{30b}$$

$$\gamma = (1/C_o)(2\varepsilon_s qN)^{1/2} \tag{30c}$$

$$\sigma = 2\varepsilon_s/(C_o LE) \tag{30d}$$

$$\nu = \pi\varepsilon_s/(C_o WE). \tag{30e}$$

Notice that the threshold voltage does not contain a VD term. This approach was taken to simplify the parameter extraction process.

2.3.6.3.2 Drain Current. The application of Ohm's law to the channel current leads to an expression for the drain current. Ohm's law for an n-MOSFET is

$$J_y = qn\mu E_y \tag{31}$$

where $n = |Q_c'|/(LE \bullet WE \bullet X_c) = |Q_c|/X_c$ where X_c is the effective channel thickness. The current density $J_y = I_c/(WE \bullet X_c)$ so that the channel current is

$$I_c = \mu WE|Q_c|E_y. \tag{32}$$

By substituting $E_y = -dV_y/dy$ and letting $I_c = -I$, the drain current is:

$$I = \mu WE|Q_c|dV_y/dy. \tag{33}$$

The channel mobility is given approximately [3, 18] by:

$$\mu = \mu_o/(1 + K_\theta E_x + |E_y|/E_c) \tag{34}$$

where μ_o is the zero-field channel mobility, E_x and E_y are the electric fields in the channel in the x- and y-directions, respectively, and E_c is a critical field parameter associated with carrier velocity saturation. The E_x field is related to Q_g by

$$\varepsilon_o E_x = Q_g. \tag{35}$$

The solution for the drain current follows by integrating the above three equations or

$$\int_0^{LE} Idy = \int_0^{VD} (\mu_o WE|Q_c| - K_\theta Q_g I \cdot LE/(\varepsilon_o VD) - I/E_c)dV_y \qquad (36)$$

where the approximation $dV_y/dy \simeq VD/LE$ was used [3].

After the integration and using Eqs. (18) and (29)

$$I = (\mu_o C_o WE/LE)(VE - aVD/2)VD - (K_\theta C_o/\varepsilon_o)(VF - VD/2)I - (1/(E_c LE))I \cdot VD \qquad (37)$$

After rearranging terms, the expression for the drain current in the linear region is:

$$I = \beta(VE - aVD/2)VD/(1 + \theta VF + \zeta VD) \qquad (38)$$

where

$$\beta = \mu_o C_o WE/LE \qquad (39a)$$

$$\theta = K_\theta C_o/\varepsilon_o \qquad (39b)$$

$$\zeta = -\theta/2 + (1/E_c)/LE \qquad (39c)$$

Equation (38) has been carefully arranged so that the VG- and VD-dependent terms appear explicitly and are not imbedded in any of the coefficients. That is, VG and VD do not appear in VT, δ, or θ. As will be shown later, ζ has a complicated dependence on VD in the linear region; thus $\zeta = \zeta(VD)$. In the saturation region where VD > VX, $\zeta = \varepsilon$. This formulation for Eq. (38) allows a precise curve fit for a given VB, W, and L. Once the transistor curves are fitted, the VB- and geometry-related coefficients are determined using the global fitting procedure. In Eq. (38), θ represents the VG-dependent mobility degradation and ζ the VD-dependent mobility degradation. Thus θ represents the mobility degradation due to E_x at the source, where E_x is a maximum. As the carriers travel toward the drain, the E_x field is reduced due to VD. This effect and the mobility degradation due to E_y are accounted for by ζ given in Eq. (39c).

2.3.6.3.3 Velocity Saturation. The expression $J = \rho_c v_c$ describes the relationship between the channel current density, J (A/cm^2), the channel volume

charge density, ρ_c (coulomb/cm^3), and the critical limiting velocity, v_c (cm/s), for the channel carriers.

Before the channel carrier velocity saturates, the velocity is given by $v = \mu E_y$, where the channel mobility is given by Eq. (34). The velocity is plotted in Figure 2.3.6-4, where it is seen that for $E_y >> E_c$ and for $K_\theta = 0$, $v_c = \mu_o E_c$. As the channel carriers travel from the source to the drain, their velocity increases due to the increase in the lateral electric field, E_y. An expression for E_y and its dependence on the channel potential, V_y, can be found by combining Eqs. (29) and (32):

$$E_y = I/(\mu WE \bullet C_o(VE - aV_y)) \tag{40}$$

Equations (29) and (40) illustrate that as aV_y approaches VE, Q_c becomes zero and E_y becomes infinitely large. But before this happens, E_y far exceeds E_c and the velocity of the channel carriers saturate.

When channel carriers first reach velocity saturation at the drain, the drain voltage is denoted VX. This voltage is traditionally known as VDsat. At this condition the drain current, derived from the convection current density, $J = \rho_c v_c$, is:

$$I = v_c WE|Q_c(VX)| \tag{41}$$

where, from Eq. (29), $|Q_c(VX)| = C_o(VE - aVX)$ and $v_c = \mu_o E_c$.

An expression for the onset of saturation was obtained by equating the above equation to the linear region drain current, Eq. (38), evaluated at VD = VX where $\zeta = \varepsilon$:

$$(\mu_o C_o WE/LE)(VE - aVX/2)VX/(1 + \theta VF + \varepsilon VX) = \mu_o E_c WE \bullet C_o(VE - aVX). \tag{42}$$

This expression differs from Hanafi's Eq. (30), where he used $v_c = \mu E_c$. The slight change from μ to μ_o in the v_c expression leads to a different VX expression and complicates the linear-saturation region slope matching, a subject not explicitly discussed by Hanafi. Note that ε is introduced for the first time in Eq. (42). It denotes the value for ζ when VD = VDsat = VX. The evaluation of Eq. (39c) at $\zeta(VX) = \varepsilon$ yields:

$$E_c LE = 1/(\varepsilon + \theta/2). \tag{43}$$

The combination of the above equations leads to an important analysis, termed the saturation equation:

$$(VE - aVX/2)VX/(1 + \theta VF + \varepsilon VX) = (VE - aVX)/(\varepsilon + \theta/2). \quad (44)$$

This equation will be used to derive the expression for VX and to aid in the matching of the drain current slopes between the linear and saturation regions.

2.3.6.3.4 VDsat = VX Derivation. The above equation is used to define the onset of the saturation region. The solution for VX requires the solution to the following equation that is quadratic in VX, that is,

$$VE(1/VX)^2 - ac(1/VX) - ad\varepsilon/(2b) = 0. \quad (45)$$

where $a = 1 + \delta$, $b = 1 + \theta VF$, $c = 1 + \theta VE/(2ab)$, and $d = 1 - \theta/(2\varepsilon)$. The solution to this equation is:

$$VX = (2VE/(ac))/[1 + (1 + 2d\varepsilon VE/(abc^2))^{1/2}] \quad (46)$$

where the plus sign was chosen for the sign of the square root term. The general form of this expression is close to the VDsat presented by Silburt et al. [2] in their Eq. (2).

2.3.6.3.5 Saturation Region Expression. In order to ensure continuity of the drain current between the linear and saturation regions, the expression for the saturation region current is equated to the linear region expression at VD = VX [3], and for VD > VX, a λ term is added to account for the slope of the drain current in the saturation region. For $VD \geq VX$ and $VG \geq VT$, the drain current in the saturation region is:

$$I = \beta(VE - aVX/2)VX/[1 + \theta VF + \varepsilon VX - \lambda(VD - VX)] \quad (47)$$

where λ is traditionally known as the channel length modulation factor. In this treatment λ is interpreted as the length of the region extending from the drain toward the source where the carriers are velocity-saturated. Various expressions for λ can be found in the literature. As discussed by Silburt et al. [2], the λ term can be located in the numerator of the saturation current expression. For the transistors analyzed in this study, the simple form for the λ expression as incorporated in Eq. (47) was found to be adequate.

2.3.6.3.6 Linear-Saturation Region Transition. To aid the stability of the network solving code (e.g., SPICE [1]), the slopes of the drain current curves were matched at the transition between the linear and saturation regions. Before taking the derivatives, the linear current from Eq. (38) is written as:

$$I = \beta(VE - aVD/2)VD/(1 + \theta VF + \mu_\zeta) \tag{48}$$

where $\mu_\zeta = \zeta VD$, $\mu_\zeta(VX) = \varepsilon VX$, and $\zeta = \zeta(VD)$:

$$dI/dVD|_{VX} = \beta(VE - aVX)/(1 + \theta VF + \varepsilon VX) - \beta(d\mu_\zeta/dVD|_{VX})(VE - aVX/2)VX/(1 + \theta VF + \varepsilon VX)^2 \tag{49}$$

The evaluation of the saturation-region drain current, Eq. (47), at VX is

$$dI/dVD|_{VX} = \lambda\beta(VE - aVX/2)VX/(1 + \theta VF + \varepsilon VX)^2 \tag{50}$$

After equating the above two equations and substituting into Eq. (44), the result is:

$$d\mu_\zeta/dVD|_{VX} = \varepsilon - \lambda + \theta/2. \tag{51}$$

2.3.6.3.7 Linear Region μ_ζ Analysis. The parameter μ_ζ was introduced in the linear current expression to account for the degradation in carrier mobility due to the lateral electric field. The evaluation of μ_ζ involves evaluating four boundary conditions in the linear current region at VD = 0 and VD = VX. They are:

1. $\mu_\zeta(0) = 0.$ (52a)

2. $d\mu_\zeta/dVD|_{VD\to 0} = \eta.$ (52b)

3. $\mu_\zeta(VX) = \varepsilon VX.$ (52c)

4. $d\mu_\zeta/dVD|_{VD\to VX} = \varepsilon - \lambda + \theta/2\ .$ (52d)

Note that η is introduced in Eq. (52b) and it denotes the slope of the μ_ζ versus VD curve at VD → 0.

A polynomial, of the form $\mu_\zeta = a + bVD + cVD^2 + dVD^3$, was used to accommodate the above four conditionsN The result can be expressed simply as $\mu_\zeta = \zeta VD$, where

$$\zeta = \eta(1 - 2V_\zeta + V_\zeta^2) + \varepsilon(2 - V_\zeta)V_\zeta + (\lambda - \theta/2)(1 - V_\zeta)V_\zeta \qquad (53)$$

where $V_\zeta = VD/VX$. As will be seen, μ_ζ is used to compensate for the difference in the β's in the linear and saturation regions.

2.3.6.4 Solution Sequence. The solution sequence was designed to reduce the number of parameters and retain optimum fitting integrity. This approach involves the evaluation of a set of transistors with different geometries. For example, if W and L are the minimum transistor channel width and length, respectively, then a suitable set of four transistors might have the following geometries: (W, L), (iW, L), (iW, jL), and (W, jL), where i and j are adjusted to sample the transistor sizes of interest. In general terms the solution sequence involves solving first for the "individual" transistor parameters and then solving for the "global" transistor parameters that account for the geometrical and VB variation in the individual transistor parameters. The individual and global parameters are listed in Table 2.3.6-1. The solution sequence is described in Table 2.3.6-2, where the subroutines BETA, TAU, THRESHOLD, LIN#1, SAT#1, SAT#2, and LIN#2 are used to derive the individual transistor parameters. It is noted in passing that certain global transistor parameters, namely LE, WE, KP, KLT, KWT, θ, and ψ, are required to evaluate the individual transistor parameters.

2.3.6.4.1 Individual Transistor Parameters. The solution for the individual transistor parameters is complicated by having to sort the transistor data between the subthreshold, linear, and saturation regions. As seen in Table 2.3.6-2, this difficulty is managed by solving the fitting algorithm a number of times. For the first few passes less than N_1, the data is sorted between the linear and saturation regions using an abbreviated version of VDsat; that is, VE/a. For subsequent passes, the data is sorted using VDsat = VX given by Eq. (46). The convergence of the algorithm depends on the spacing between the data in the data set. A convergence criteria or the number of passes through the algorithm can be specified. In this effort $N_1 = 3$ and $N_2 = 5$. To reiterate, the number of passes through the algorithm can only be specified with a limited accuracy, for any change in the boundary between the regions alters the data set and hence the values of the parameters.

2.3.6.4.1.1 LIN#1 Subroutine. This subroutine is used to determine VT, ζ, and τ from the solution to the extrinsic drain current in the linear region.

As seen in Figure 2.3.6-5, the extrinsic voltages, UG and UD, are related to the intrinsic voltages, VG and VD, or:

$$UG = VG + IR \tag{54}$$

$$UD = VD + 2IR \tag{55}$$

The derivation of the extrinsic drain current in the linear region follows from the combination of Eqs. (38) and (55):

$$UD/I = 2R + (1 + \theta VF + \zeta VD)/\beta(VE - aVD/2) \tag{56}$$

This can be reduced to

$$I/UD = \beta(VE - aVD/2)/[(1 - 2\beta RVT - \theta\psi) + (\theta + 2\beta R)VG + (\zeta - \beta R\delta - \beta R)VD] \tag{57}$$

By substituting the expression for ζ, Eq. (53), and defining

$$\alpha = 1 - 2\beta RVT - \theta\psi \tag{58a}$$

$$\alpha 1 = -2\eta + 2\varepsilon + \lambda - \theta/2 \tag{58b}$$

$$\alpha 2 = \eta - \varepsilon - \lambda + \theta/2 \tag{58c}$$

$$\tau = \theta + 2\beta R \tag{58d}$$

$$\rho = \eta - \beta R\delta - \beta R \tag{58e}$$

the following equation results:

$$I/UD = [\beta/\alpha](VG - VD/2) - [\beta VT/\alpha]1 - [\beta\delta/\alpha]VD/2 - [\tau/\alpha]I\bullet VG/UD - [\rho/\alpha]I\bullet VD/UD - [\alpha 1/\alpha]I\bullet VD^2/(UD\bullet VX) - [\alpha 2/\alpha]I\bullet VD^3/(UD\bullet VX^2). \tag{59}$$

The coefficients for this linear equation are identified by the bracket terms. The coefficients are: C1 = β/α, C2 = $-\beta VT/\alpha$, C3 = $-\beta\delta/\alpha$, C4 = $-\tau/\alpha$, C5 = $-\rho/\alpha$, C6 = $-\alpha_1/\alpha$, and C7 = $-\alpha 2/\alpha$. From these coefficients, the following individual transistor parameters are determined: VT = -C2/C1, δ = -C3/C1, and τ = $-\alpha$C4. Also, the linear region β is determined from $\beta = \alpha C1$ in order to compare its value with the saturation region β. Notice that VT and δ are determined without a knowledge of α, but the calculation of τ and β requires α (see Eq. (58a)). The parameters involved in α will be determined subsequently (see Eqs. (69) and (70)). The fitting procedure uses the least-squares method with a Gauss-Jordan matrix inversion routine [10].

2.3.6.4.1.2 SAT#1 Subroutine. This subroutine is used to determine β, ε, and λ from the intrinsic saturation drain current, Eq. (47). The β(VB = 0) values are used to determine the ΔW and ΔL values; these values are used subsequently to determine certain global transistor parameters. The equation is:

$$I(1 + \theta VF) = [\beta](VE - aVX/2)VX - [\varepsilon]I \cdot VX + [\lambda](VD - VX)I. \qquad (60)$$

This equation must be solved in concert with the VX expression given by Eq. (46), which requires an iterative solution based on the parameter ε.

2.3.6.4.1.3 SAT#2 Subroutine. This subroutine is used to reevaluate ε values for VB < 0 drain curves using the intrinsic saturation drain current, Eq. (47). In this solution β(VB < 0) values are equated to β(VB = 0) values as determined from SAT#1. This adjustment in β(VB < 0) values is compensated for by adjusting ε(VB < 0) values. The equation is

$$I(1 + \theta VF - \lambda(VD-VX)) - \beta(VE - aVX/2)VX = -[\varepsilon]VX \cdot I. \qquad (61)$$

This equation must be solved along with VX, Eq. (46), which requires an iterative solution based on ε.

2.3.6.4.1.4 LIN#2 Subroutine. This subroutine is used to determine η using the intrinsic drain current Eq. (38). In this solution β values are equated to β(VB = 0) values determined in SAT#1. This adjustment in the β values is compensated for in the η value. The combination of Eqs. (38) and (53) leads to:

$$(I/VD)(1 + \theta VF) - \beta(VE - aVD/2) + \varepsilon(2 - V_\zeta)V_\zeta I$$
$$+ (\lambda - \theta/2)(1 - V_\zeta)V_\zeta I = -[\eta](1 - 2V_\zeta + V_\zeta)I \qquad (62)$$

where V_ζ = VD/VX.

2.3.6.4.2 Global Transistor Parameters. The global transistor parameters are a set of 19 parameters determined from the seven individual transistor parameters measured on a number of transistors with different L and W geometries. The algebraic form for the global transistor equations follows from the theoretical equations derived in Section 2.3.6.3. The global parameters contain the W, L, and VB dependence of the individual transistor parameters. As shown below, three additional parameters are derived from the global transistor parameters. These are VTO, R, and RW.

2.3.6.4.2.1 BETA Subroutine. This subroutine is used to determine KP, LE, and WE from Eq. (5) given W, L, and β. The linearization of Eq. (5) leads to

$$\beta L = [KP]W - [KP \cdot \Delta W]1 + [\Delta L]\beta \tag{63}$$

where the coefficients are C1 = KP, C2 = $-KP \cdot \Delta W$, and C3 = ΔL, so the parameters are KP = C1, ΔW = C2/C1, and ΔL = C3. From these coefficients one can derive the effective channel length LE = L - ΔL and the effective channel width WE = W - ΔW.

2.3.6.4.2.2 THRESHOLD Subroutine. This subroutine is used to determine ψ, γ, KLG, and KWG. The equation is derived by combining Eqs. (30a), (30d) and (30e); that is,

$$VT = [\psi]1 + [\gamma](\phi - VB)^{1/2} - [KLG](\phi - VB)/LE + [KWG](\phi - VB)/WE \tag{64}$$

where $KLG = 2\varepsilon_s/C_o$ and $KWG = \pi\varepsilon_s/C_o$. In this solution the value for ϕ is taken as 0.6 volts. The choice for ϕ does not influence the VB $\neq$ 0 curve fitting and was found to have a negligible influence on the VB = 0 curve fitting. It is customary [19] to calculate the zero-bias (VB = 0) threshold voltage for large geometry transistors from:

$$VTO = \psi + \gamma(\phi)^{1/2} \tag{65}$$

2.3.6.4.2.3 DELTA Subroutine. This subroutine is used to determine DO, KLD, and KWD. The equation follows from the combination of Eqs. (30b), (30d), and (30e), or:

$$\delta = [DO]/(\phi - VB)^{1/2} - [KLD]/LE + [KWD]/WE \tag{66}$$

where DO = $\gamma/2$, $KLD = 2\varepsilon_s/C_o$ and $KWD = \pi\varepsilon_s/C_o$. Notice that KLD = KLG and KWD = KWG. In order to give the curve-fitting routine more degrees of freedom, we decided not to equate these coefficients.

2.3.6.4.2.4 TAU Subroutine. This subroutine is used to evaluate KLT, θ, and R. The form for τ follows from Eq. (58d), which requires an evaluation of $2\beta R$. The series resistance is expressed as:

$$R = R_s L_e / WE \tag{67}$$

where R_s is the sheet resistance of the source or drain and L_e is the effective length of the series resistor. The combination of the above equation with the β from Eq. (5) leads to:

$$2\beta R = 2KP \cdot R_s L_e / LE. \tag{68}$$

Substituting this expression for $2\beta R$ into Eq. (58d) leads to:

$$\tau = [\theta]1 + [KLT]/LE \tag{69}$$

where

$$KLT = 2KP \cdot R_s L_e \tag{70}$$

From the combination of Eqs. (67) and (70), the series resistance is

$$R = KLT/(2KP \cdot WE) \tag{71a}$$

and the series resistance coefficient is

$$RW = KLT/(2KP) \tag{71b}$$

2.3.6.4.2.5 ETA Subroutine. This subroutine is used to evaluate HO and KLH. The form for η follows from the ζ expression, Eq. (39c), or:

$$\eta = [HO]1 + [KLH]/LE. \tag{72}$$

From Eq. (39c), the sign for HO should be negative, but in practice the sign is positive.

2.3.6.4.2.6 EPSILON Subroutine. This subroutine is used to evaluate EO and KLE. Recall that $\zeta(VX) = \varepsilon$, so the form for ε follows from the ζ equation, Eq. (39c), or

$$\varepsilon = -[EO]1 + [KLE]/LE \tag{73}$$

where $EO = \theta/2$ and $KLE = 1/E_c$.

2.3.6.4.2.7 LAMBDA Subroutine. This subroutine is used to evaluate LO, KLL, and KWL. The form for λ is taken as:

$$\lambda = [LO]1 + [KLL]/LE - [KWL]/WE. \tag{74}$$

2.3.6.4.3 Parameter Relationships. The geometrical dependence for the seven individual transistor parameters was derived in some cases from the theoretical analysis but in all cases was taken to have a linear dependence on 1/WE and 1/LE. Such a dependence is satisfactory for small variations in LE and WE. For large variations in LE and WE, the geometrical dependence may have to be altered. For instance, in CASMOS [6], the eta parameter was found to have a $1/LE^2$ dependence and the lambda parameter was found to have a truncated geometry dependence. To accommodate more complex geometrical dependences, higher order terms could be included in the above global parameter expressions.

From the theory and fitting equations, certain relationships are expected among the parameters; that is, $DO = \gamma/2$, KLD = KLG, and KWD = KWG. One is tempted to reduce the number of coefficients by using the above equalities. But the above coefficients were retained to allow more flexibility in the analysis. As shown in the following results, the equalities KLD = KLG and KWD = KWG do not hold. This is to be expected since these coefficients represent the first two terms in the series expansion of the term that accounts for the body effect.

2.3.6.5 Results. The experimental results presented here were taken from a number of 1.2-μm and 3-μm CMOS transistors. First we will discuss results from 1.2-μm n-channel transistors with geometries indicated in Figure 2.3.6-6. The data for each transistor consists of a set of drain currents measured for VD = 0.05, 0.2, 0.6, 1.0, 2.0, 3.0, 4.0 and 5.0 volts, and VG = 0.5 to 5.0 volts in 0.5-volt increments with $|VB| = 0$ and 2.5 volts. A BASIC program, JMOSFIT, was written which performs the multiple-parameter least-squares fitting. The program runs on a VAX 11/780 computer. For the above data set, the analysis requires less than 1 minute of CPU time to run a precompiled program.

2.3.6.5.1 1.2-μm n-MOSFET Individual Transistor Fit Analysis. Intermediate results from the fitting of four 1.2-μm n-channel transistors are listed in Table 2.3.6-3. These results represent the output from the four subroutines LIN#1, SAT#1, SAT#2, and LIN#2. By examining the results in Table 2.3.6-3, one can visualize the working of each subroutine by observing the input (GIVEN) parameters and output (FIND) parameters. The results from the four fitting subroutines are listed in Table 2.3.6-3.

The transistors are arranged in a geometrical order such that the first three transistors are expected to have increasing threshold voltages according to classical short and narrow channel effects. Referring to Figure 2.3.6-6, the threshold voltage of transistor #1 is degraded due to short channel effects, the threshold voltage of transistor #2 is a reference value for a "large" channel device, the threshold voltage of transistor #3 is enhanced by narrow channel effects, and the threshold voltage of transistor #4 results from geometry fighting of narrow and short channel effects. As discussed in Section 2.3.6.5.4, the threshold voltage may not follow the expectation of the classical short and narrow channel effects.

A summary of the individual transistor fitting results is shown at the top of Table 2.3.6-4. These parameters were used to compute the drain curves shown in Figures 2.3.6-7 and 2.3.6-8. The computed points were compared with the original experimental data points and a correlation coefficient (cc) [10] computed. As seen in Table 2.3.6-4, the correlation coefficient is better than three nines for all curves and is four nines in two cases.

2.3.6.5.2 1.2-μm n-MOSFET Global Transistor Parameters. The global transistor parameters are listed in the middle of Table 2.3.6-4. These values were obtained from a least-squares fit of the equations given in Section 2.3.6.4.2 using the individual transistor parameters given at the top of the table. These parameters will be discussed in Section 2.3.6.5.4.

2.3.6.5.3 1.2-μm n-MOSFET Global Transistor Fit Analysis. The analysis is shown at the bottom of Table 2.3.6-4. From the values shown in this table, one can verify that the analysis method has been implemented correctly. For example, the BETA values for VB < 0 are equal to $\beta(VB = 0)$. By comparing the global transistor parameters at the bottom of Table 2.3.6-4 with the individual transistor parameters shown at the top of the table, one can assess the magnitude of the discrepancies between the global and individual transistor fits.

The global transistor parameters shown at the bottom of Table 2.3.6-4 were used to compute drain curves shown in Figures 2.3.6-9 and 2.3.6-10. The computed points were compared with the original experimental data points and a correlation coefficient computed. As seen in Table 2.3.6-4, the correlation coefficient is better than two nines for all curves.

The global fits could be improved in several ways. The global fitting equations could have a more complex geometry and VB dependence. Or the region of the width-length plane sampled could be reduced so that the linear global fitting equations are adequate.

2.3.6.5.4 Additional Results. Results from 1.2-μm p-channel MOSFETs are listed in Table 2.3.6-5, from 3-μm n-channel MOSFETs in Table 2.3.6-6, and from 3-μm p-channel MOSFETs in Table 2.3.6-7. The global fitting, as judged by the correlation coefficient, is comparable in all cases. From the classical theory presented in Section 2.3.6.3.1, VT is expected to decrease for shorter transistors and to increase for narrower transistors. But, as seen in Table 2.3.6-5 for the 1.2-μm p-MOSFET, the threshold voltage for VB = 0 does not follow the classical short and narrow channel predictions. For VB = 2.5 V, VT does follow the classical predictions for the shorter devices, but for narrower devices the VT variation remains nonclassical. This has resulted in the negative sign for the KWG parameter listed in Table 2.3.6-5. This behavior has been observed by others [20, 21] and is attributed to perturbations in the electric fields due to nonuniform dopant densities in the bulk, doping of the field channel stop, and the nature of the field-to-gate-oxide transition.

A summary of results from a number of transistors is listed in Tables 2.3.6-8 and 2.3.6-9. The strength of the variation of VT with geometry is given by the KLG and the KWG parameters. For these transistors these parameters are relatively small, which means that VT is not a strong function of geometry. Recall from the theory that we purposely did not set KLG = KLD and KWG = KWD, and, as seen in the tables, this was a prudent choice, for DELTA has a much larger geometry dependence than VT.

From the analysis of DELTA in Eq. (66), it is expected that DO = GAMMA/2 and, as seen in the tables, this trend is followed approximately. That is, the DO values are within 33 percent of GAMMA/2.

The THETA values appear to be independent of dimensions of the technology and depend only on the channel type. From Table 2.3.6-8, the value of n-THETA is 0.037 ± 0.003 (1/V) for n-MOSFETs, and from Table 2.3.6-9, p-THETA is 0.098 ± 0.001 (1/V) for p-MOSFETs. Similar values for n-MOSFETs can be found in the literature [3, 5, 22]. Values for p-THETA were not found in the literature and are perhaps reported here for the first time.

From Eq. (73), the saturated carrier velocity, critical electric-field E_c is equal to 1/KLE. From the tables the KLE values appear to be independent of the dimension of the technology and depend only on the channel type. From Table 2.3.6-8, n-E_c = 16,600 ± 1,800 V/cm, and from Table 2.3.6-9, p-E_c = 57,300 ± 8,400 V/cm. From the theory presented in Section 2.3.6.3.3, the E_c should correspond to $v_c/2$ (see Figure 2.3.6-4). But the above E_c values are within 90 percent of v_c for the early data given in the literature [23]. However, the n-E_c values are in better agreement with more recent data [18]. The exact interpretation of KLE remains under study; however, the observation that KLE is independent of the dimensions of the technology and dependent on the channel type stands.

The strength of some of the parameters can be judged by examining the tables. For example, the LO-values are very small. It may be prudent to set LO to zero and thereby reduce the number of parameters.

2.3.6.6 Conclusions. A simple MOSFET parameter extraction procedure was developed. The fitting uses seven individual transistor parameters and 19

global transistor parameters. The fitting of the individual transistor curves was achieved with a correlation coefficient of better than three nines. Four transistors were fitted that formed a rectangle in the channel width-length plane. The global fitting of these transistors was achieved with a correlation coefficient of better than two nines. A number of 1.2-μm and 3-μm CMOS transistors were analyzed, and the gate-field mobility degradation factor THETA was found to be 0.037 ± 0.003 (1/V) for n-MOSFETs and 0.098 ± 0.001 (1/V) for p-MOSFETs. The saturated carrier velocity, critical electric field E_c was found to be 16,600 ± 1,800 V/cm for n-MOSFETs and 57,300 ± 8,400 V/cm for p-MOSFETs.

2.3.6.7 Acknowledgments. The discussions with C. L. Seitz and S. Mattisson of the California Institute of Technology were particularly useful with regard to the CASMOS approach. The insight of Professor R. Winton of Mississippi State University is greatly appreciated and his derivation of the integrable form of the drain current as typified by the steps involving Eq. (36) is gratefully acknowledged. The assistance of the MOS Implementation Service, Information Sciences Institute, University of Southern California, in supplying the test transistors is greatly appreciated.

2.3.6.8 References.

1. L. W. Nagel, "SPICE2, a Computer Program to Simulate Semiconductor Circuits," Electronics Research Laboratory, University of California, Berkeley, ERL Memo No. ERL-M520 (May 1975).

2. A. L. Silburt, R. C. Foss, and W. F. Petrie, "An Efficient MOS Transistor Model for Computer-Aided Design," IEEE Trans. Computer-Aided Design, CAD-3, 104-111 (1984).

3. H. I. Hanafi, L. H. Cammitz, and A. J. Dally, "An Accurate and Simple MOSFET Model for Computer-Aided Design," IEEE J. Solid-State Circuits, SC-17, 882-891 (1982).

4. P. Yang and P. K. Chatterjee, "SPICE Modeling for Small Geometry MOSFET Circuits," IEEE Trans. Computer-Aided Design, CAD-1, 169-182 (1982).

5. F. M. Klaassen, "A MOS Model for Computer-Aided Design," Philips Res. Rep., 31, 71-83 (1976).

6. R. E. Oakley and R. J. Hocking, "CASMOS--An Accurate MOS Model with Geometry-Dependent Parameters: I," IEE Proceedings, 128, 239-247 (1981).

7. B. S. Messenger, "A Fully Automated MOS Device Characterization System for Process-Oriented Integrated Circuit Design," Electronics Research Laboratory, University of California, Berkeley, Memo No. UCB/ERL M84/18 (January 1984).

8. B. S. Sheu, D. L. Scharfetter, and H. C. Poon, "Compact Short Channel IGFET Model (CSIM)," Electronics Research Laboratory, University of California, Berkeley, Memo No. UCB/ERL M84/20 (March 1984).

9. D. E. Ward, "Optimized Extraction of MOS Model Parameters," IEEE Trans. Computer-Aided Design, CAD-1, 163-168 (1982).

10. A. R. Miller, BASIC Program for Scientists and Engineers, Sybex, Berkeley (1981).

11. S. L. Meyer, Data Analysis for Scientists and Engineers, John Wiley and Sons, New York (1975).

12. P. Yang and P. K. Chatterjee, "An Optimal Parameter Extraction Program for MOSFET Models," IEEE Trans. Electron Devices, ED-30, 1214-1219 (1983).

13. S. M. Sze, Physics of Semiconductor Devices, John Wiley and Sons, New York (1981).

14. K. F. Galloway, M. Gaitan, and T. J. Russell, "A Simple Model for Separating Interface and Oxide Charge Effects in MOS Device Characteristics," IEEE Trans. Nucl. Sci., NS-31, 1497-1501 (1984).

15. G. W. Taylor, "Subthreshold Conduction in MOSFETs," IEEE Trans. Electron Devices, ED-25, 337-350 (1978).

16. L. D. Yau, "A Simple Theory to Predict the Threshold Voltage of Short Channel IGFETs," Solid State Electron, 17, 1059-1063 (1974).

17. L. Akers, "Small-Geometry Effects," Chapter 6 in Modern MOS Technology: Processes, Devices, and Design, D. G. Ong, McGraw-Hill Book Company, New York (1984).

18. R. T. Jerdonic and W. R. Bandy, "Velocity Saturation Effects in n-Channel Deep-Depletion SOS/MOSFETs," IEEE Trans. Electron Devices, ED-25, 894-898 (1978).

19. A. Vladinirescu and S. Liu, "The Simulation of MOS Integrated Circuits Using SPICE2," Electronics Research Laboratory, University of California, Berkeley, Memo No. UCB/ERL M80/7 (October 1980).

20. M. Nishida and H. Onodera, "An Anomalous Increase of Threshold Voltages with Shortening the Channel Lengths for Deeply Boron-Implanted n-Channel MOSFETs," IEEE Trans. Electron Devices, ED-28, 1101-1103 (1981).

21. M. Sugino, L. A. Akers, and J. M. Ford, "Optimum p-Channel Isolation Structure for CMOS," IEEE Trans. Electron Devices, ED-31, 1823-1829 (1984).

22. M-S. Lin, "The Classical Versus the Quantum Mechanical Model of Mobility Degradation Due to the Gate Field in MOSFET Inversion Layers," IEEE Trans. Electron Devices, ED-32, 700-710 (1985).

23. C. B. Norris and J. F. Gibbons, "Measurement of High-Field Carrier Drift Velocities in Silicon by the Time-of-Flight Technique," IEEE Trans. Electron Devices, ED-14, 38-43 (1967).

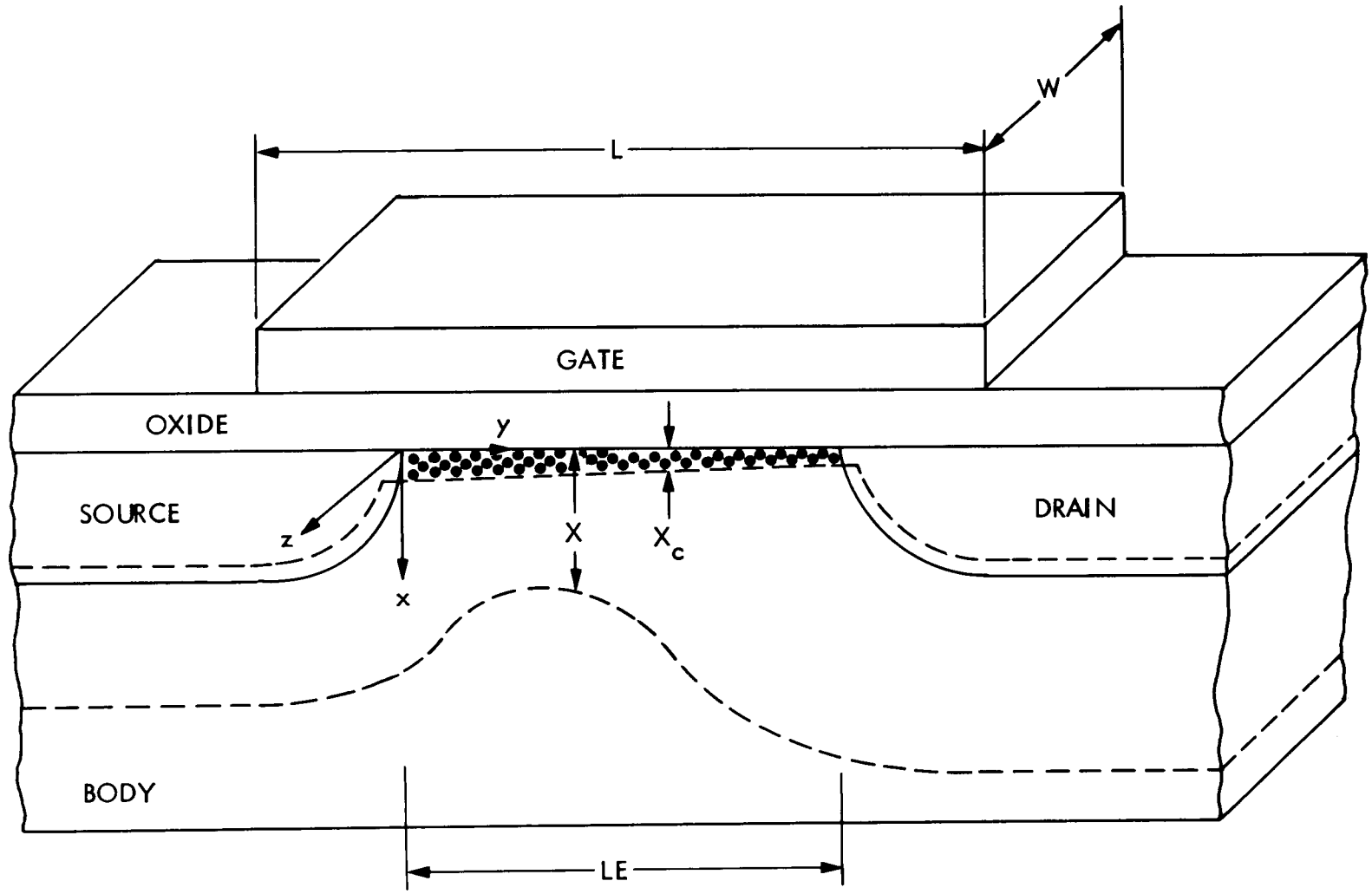


Figure 2.3.6-1. MOSFET dimensional definitions

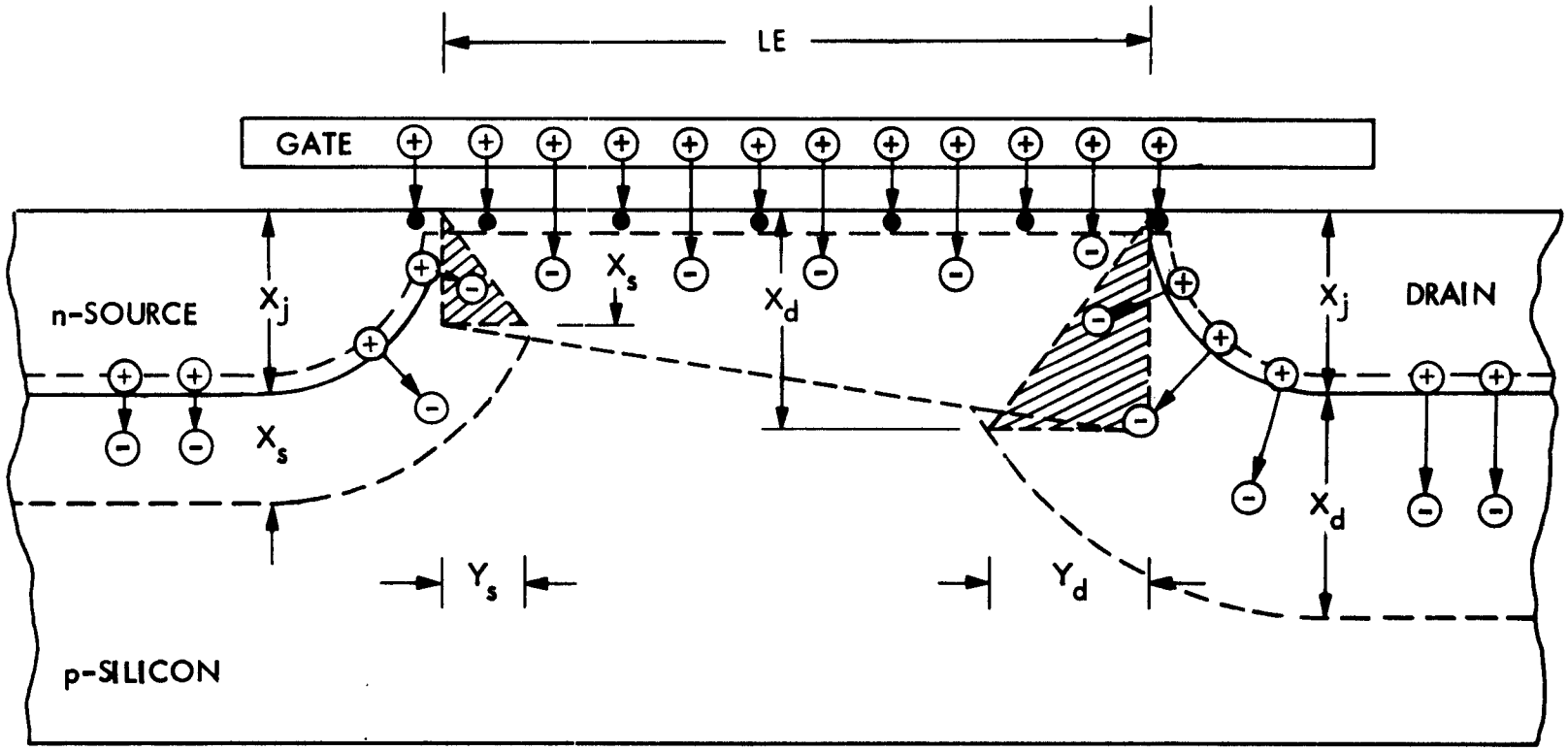


Figure 2.3.6-2. MOSFET short-channel effect due to source and drain bulk charge stealing (cross-hatched regions)

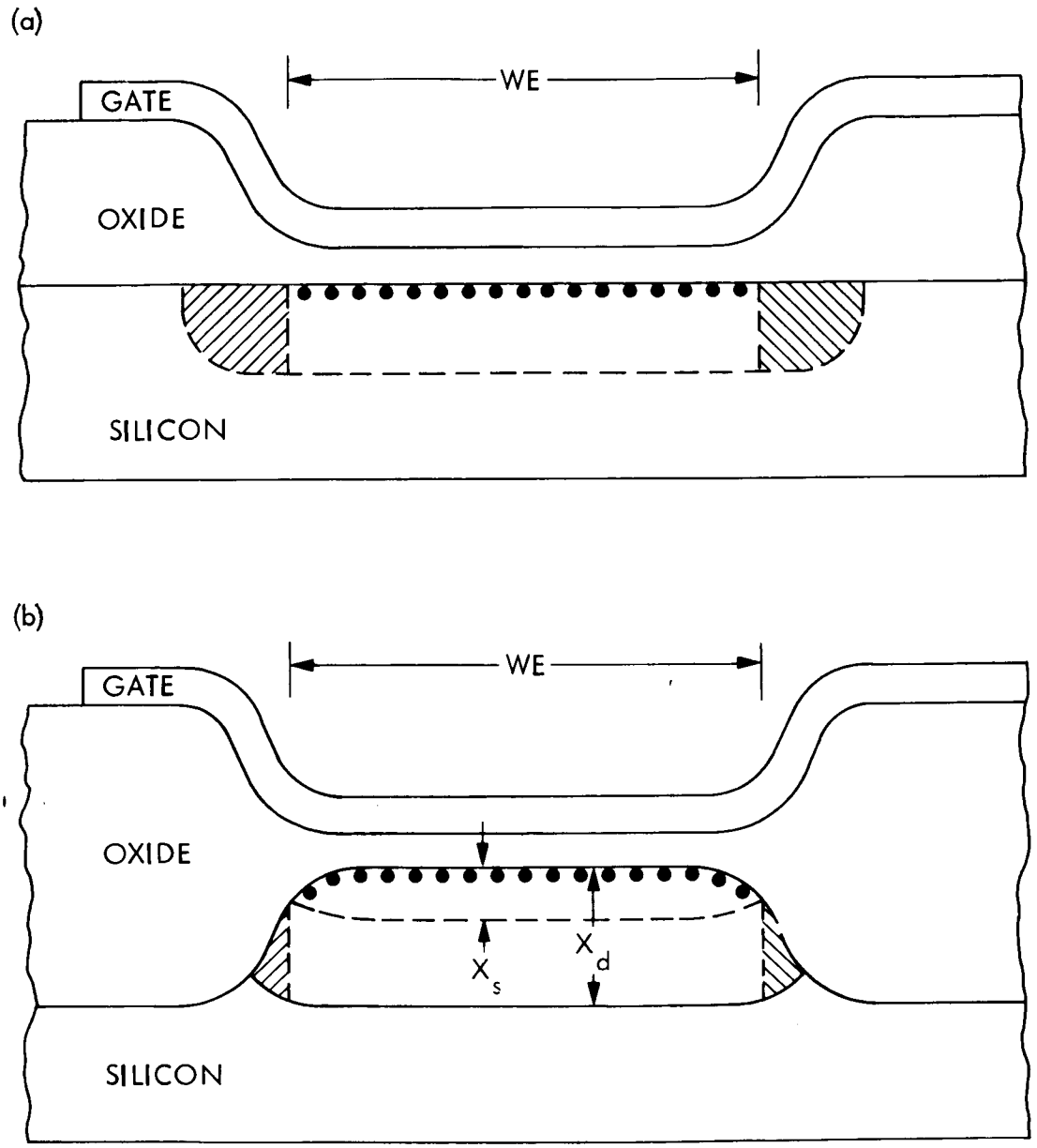


Figure 2.3.6-3. MOSFET narrow-channel effect where additional bulk charge (cross-hatched regions) appears beneath the gate with: (a) the conventional oxide process, and (b) the recessed oxide process

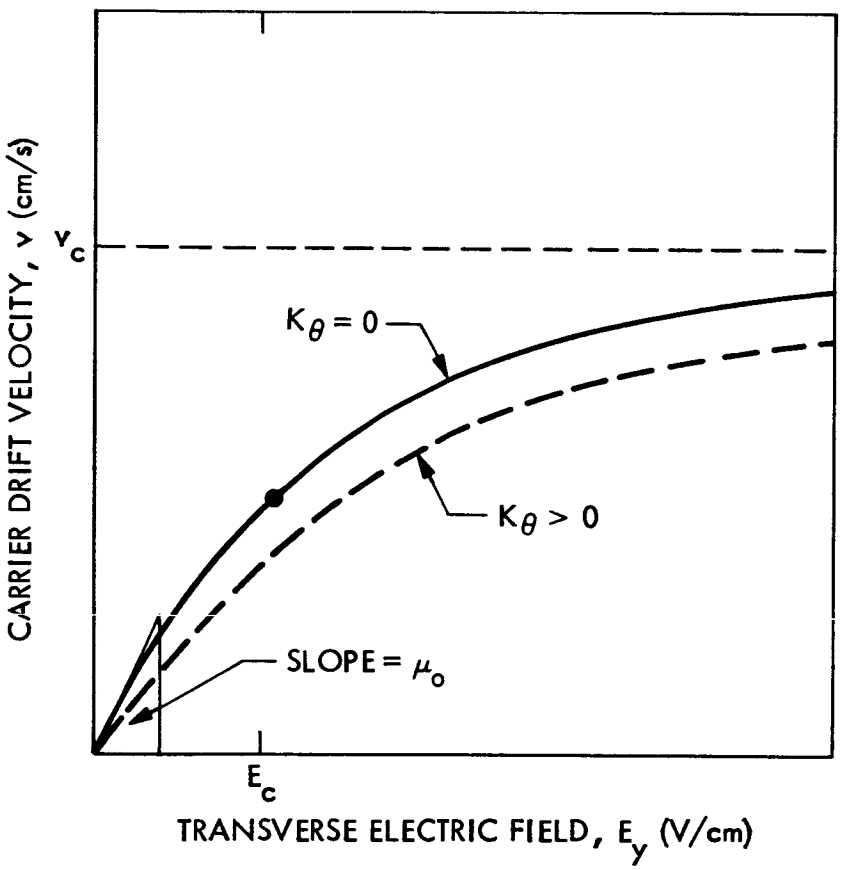


Figure 2.3.6-4. Graphical representation of the channel carrier velocity dependence on electric field

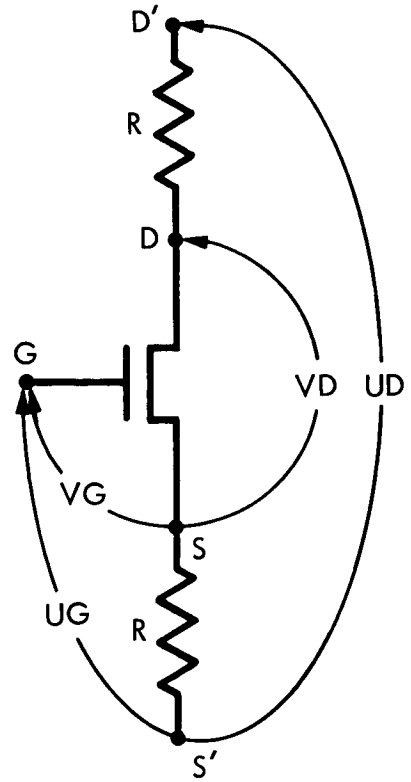


Figure 2.3.6-5. MOSFET with source and drain series resistance. The intrinsic voltages are VG and VD, and the extrinsic voltages are UG and UD.

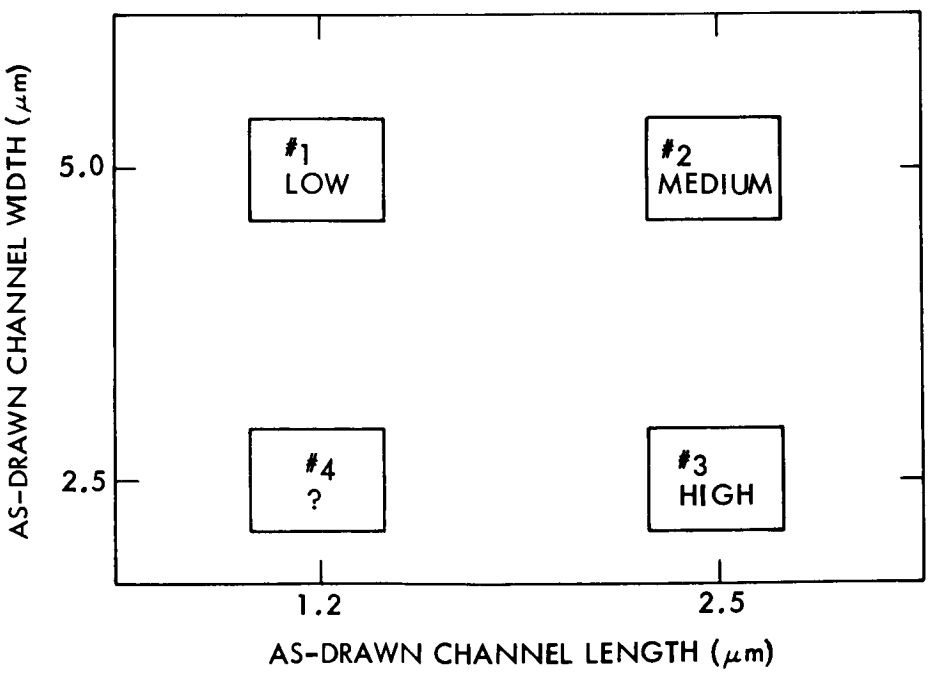


Figure 2.3.6-6. The transistor channel width-length plane where the relative threshold voltage values expected from classical short and narrow channel effects are indicated within the boxes for the four transistors used in the 1.2-μm parameter extraction

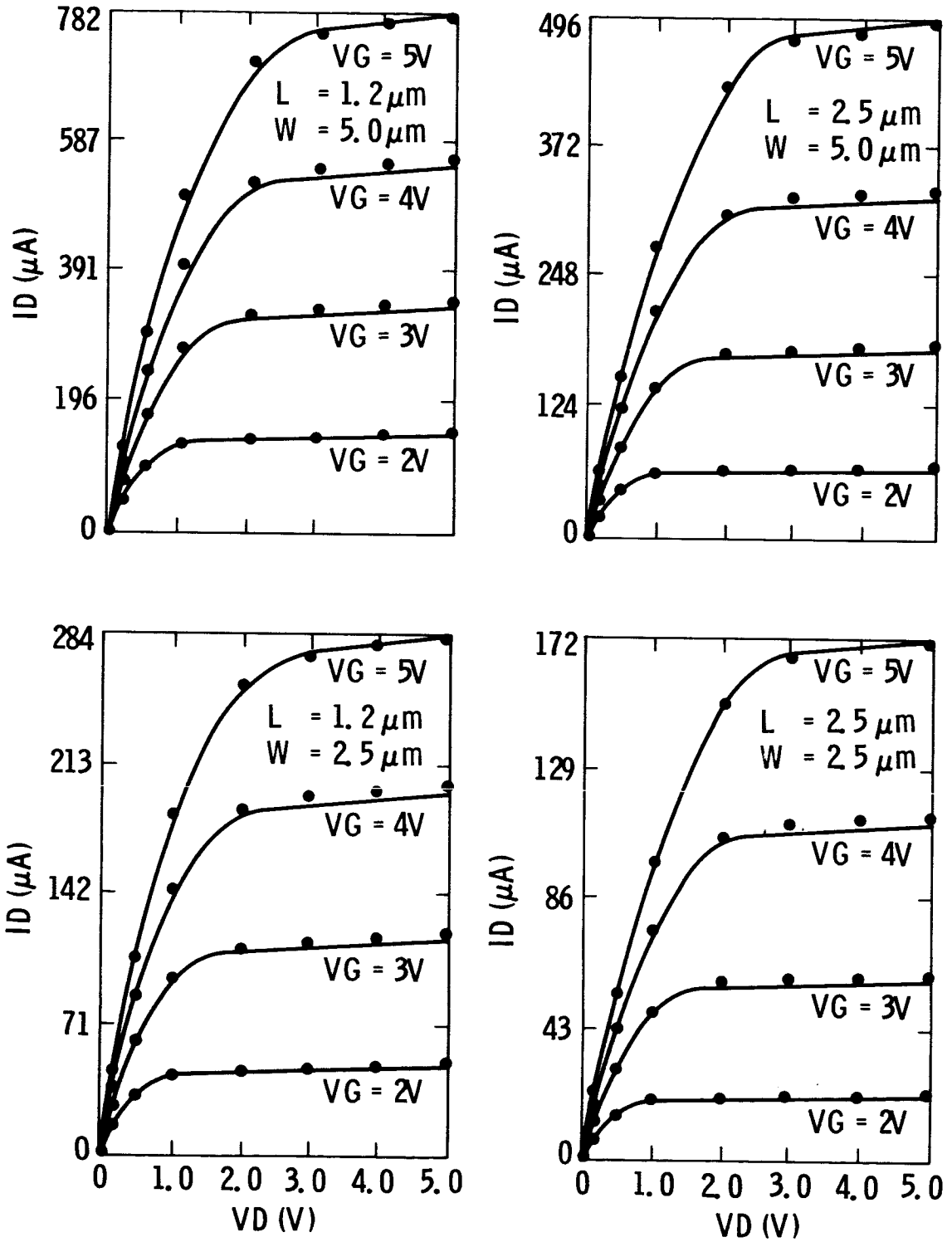


Figure 2.3.6-7. Individual fitting of four 1.2-μm CMOS n-channel transistor drain current curves for VB = 0

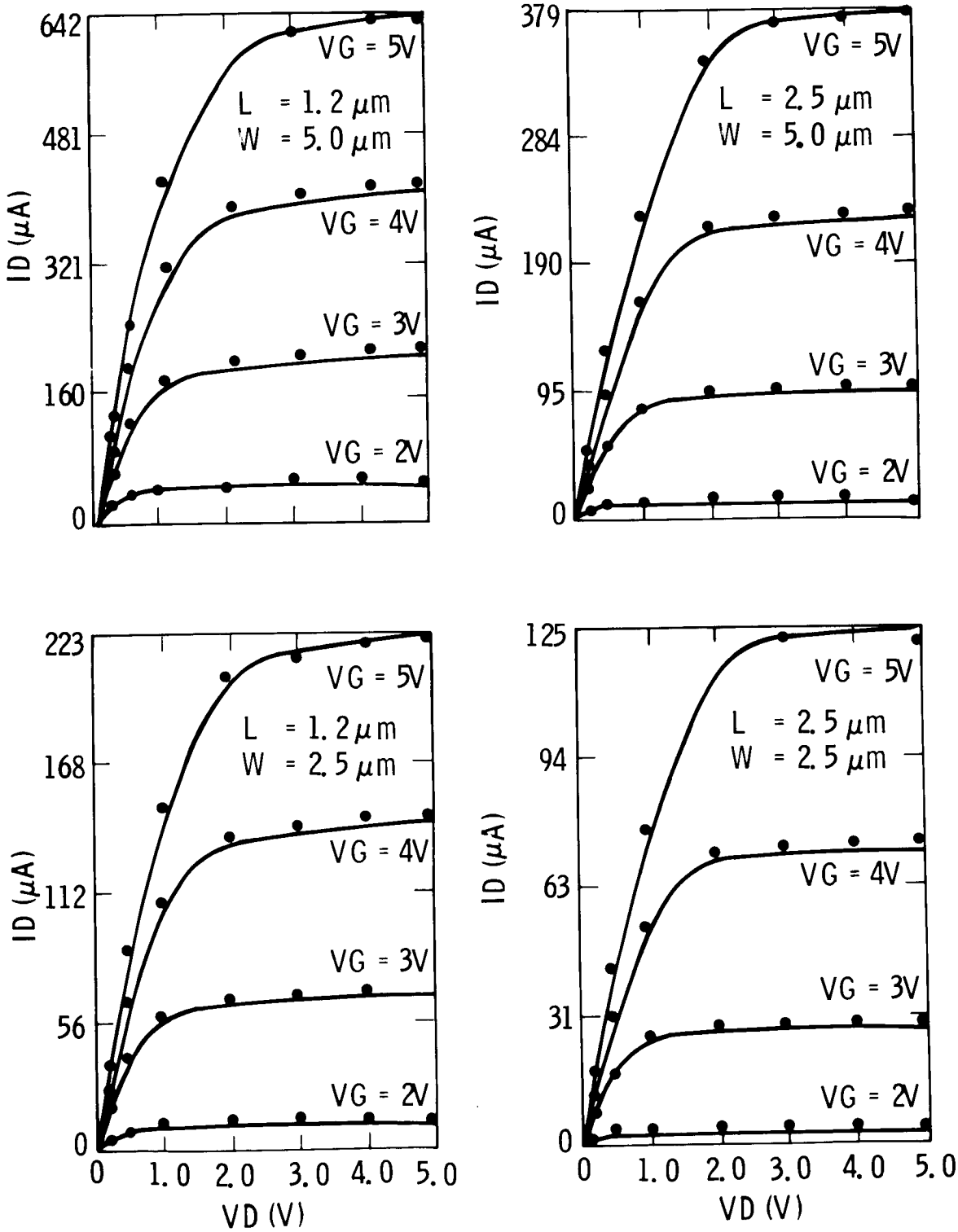


Figure 2.3.6-8. Individual fitting of four 1.2-μm CMOS n-channel transistor drain current curves for VB = -2.5 volts

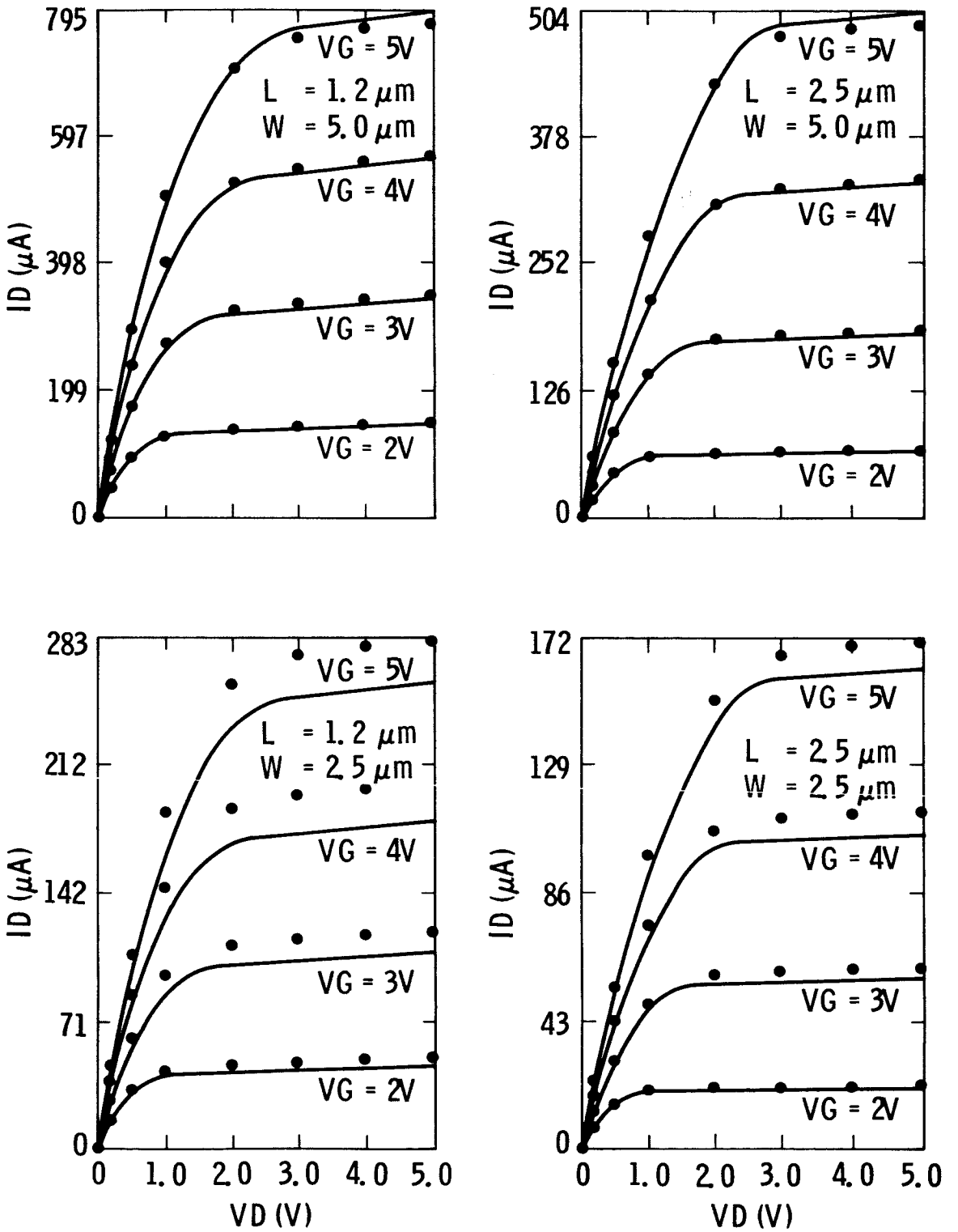


Figure 2.3.6-9. Global fitting of four 1.2-μm CMOS n-channel transistor drain current curves for VB = 0

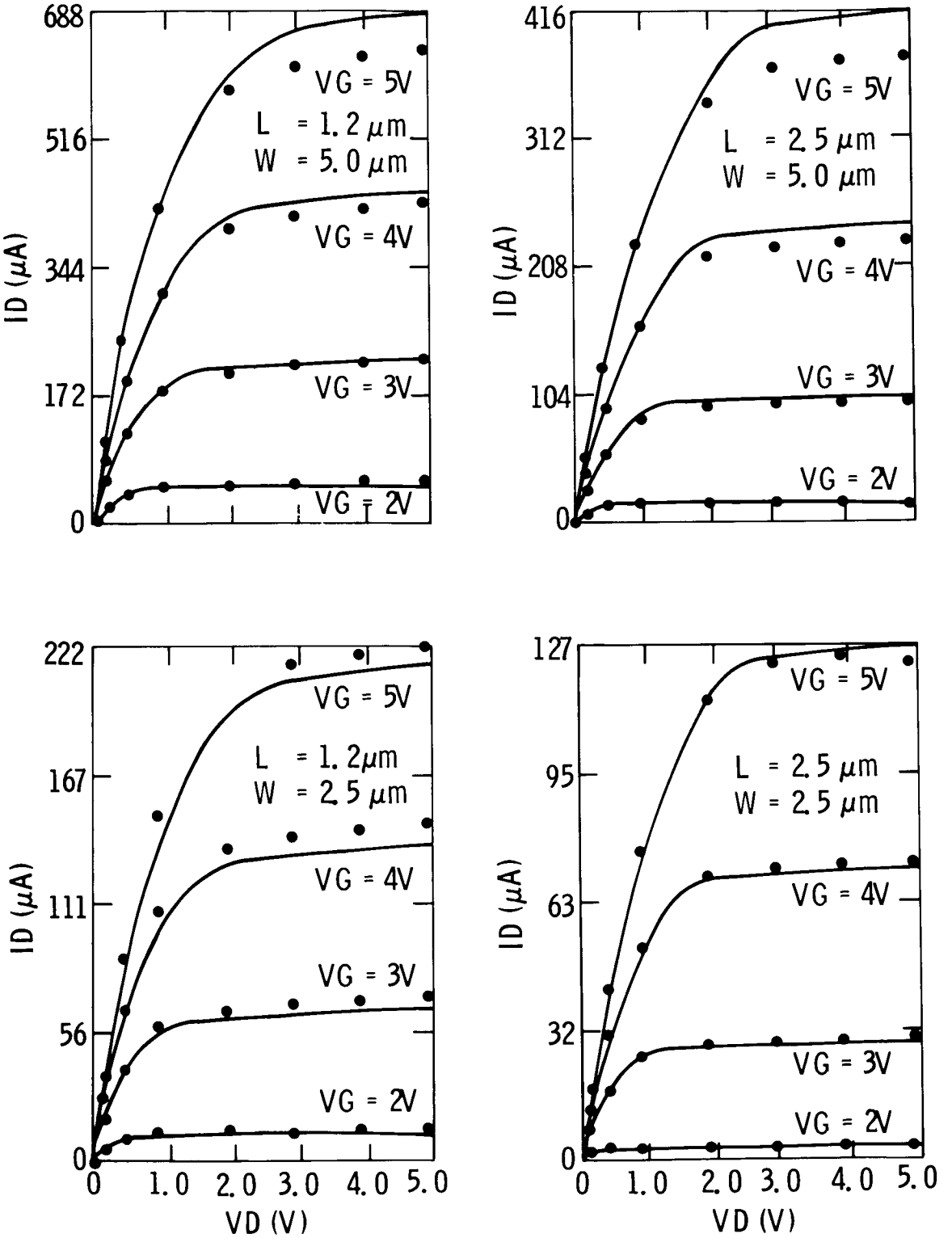


Figure 2.3.6-10. Global fitting of four 1.2-μm CMOS n-channel transistor drain current curves for VB = -2.5 volts

Table 2.3.6-1. Transistor Parameter Set

GIVEN	INDIVIDUAL	GLOBAL	DERIVED
CHANNEL LENGTH, L	BETA, β	KP, LE, WE	VTO
CHANNEL WIDTH, W	THRESHOLD, VT	ψ, γ, KLG, KWG	R
PHI, ϕ	DELTA, δ	DO, KLD, KWD	RW
	ETA, η	HO, KLH	
	TAU, τ	θ, KLT	
	EPSILON, ε	EO, KLE	
	LAMBDA, λ	LO, KLL, KWL	

Table 2.3.6-2. Parameter Extraction Solution Sequence

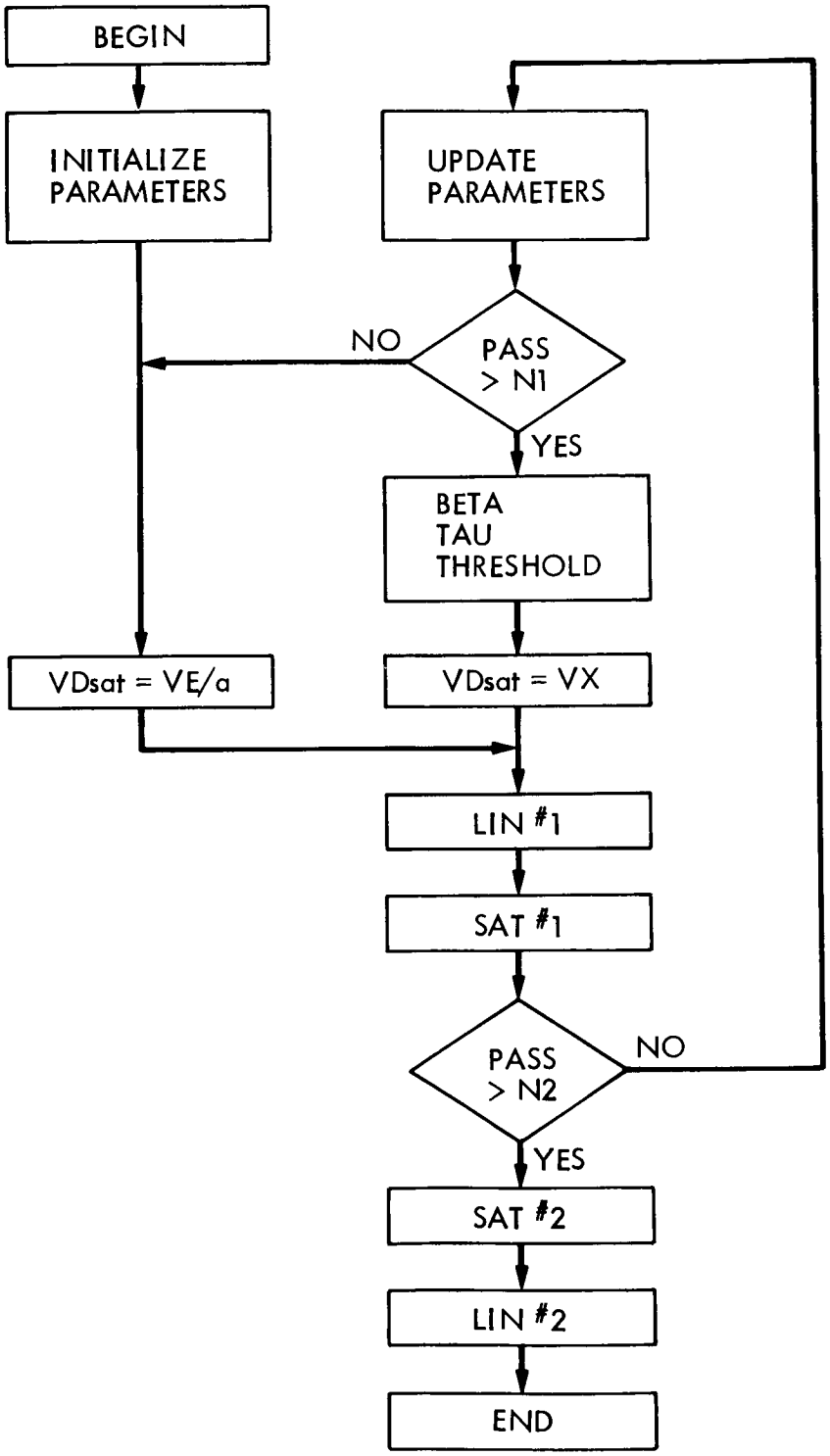


Table 2.3.6-3. Intermediate Output of JMOSFIT for 1.2-μm n-MOSFET (CMOS1N1)

```
FIT81.BAS          JMOSFIT          M. G. BUEHLER & B. T. MOORE (5-17-85)
LIN AND SAT REGION FIT TO TRANSISTOR DRAIN CURVES
ILIN = (B*(VE-A*VD/2)*VD)/(1+T*VF+H*VD)
ISAT = (B*(VE-A*VX/2)*VX)/(1+T*VF+E*VX-L*(VD-VX))
  VX = (2*VE/A)/(1+SQR(1+2*E*VE/(A*(1+T*VF)))
  VT = PSI+G*SQR(P5-V3)-KLG*(P5-VB)/LE+KWG*(P5-VB)/WE
   A = 1+D, VE = VG-VT, VF = VG-PSI
ENTER DATA FILE NAME OR RETURN FOR TEMP.TEM? CMOS1N1.DAT
ENTER THE DATA STEP SIZE? 10
RK3:NT0953.203 CREATED: 15:34:51 11-JUN-83 BY CRUNCH VERSION  6
CMOS125 N CHANNEL TRANSISTOR INPUT CHARACTERISTICS  (WAFER #9  LOC 5
N-CHANNEL TRANSISTOR INPUT CURVES OF 101 ELEMENTS EACH
PASS NO.= 5   PHI(V)= .6 PSI(V)= .100833
LIN#1 GIVEN RL AND THETAL, FIND VTL, DELTAL, TAUL AND BETAL
K%   L    W    VB    BETA     VT   DELTA  TAU                             THETA   R    DATA
 1  1.2  5.0   0.0 2.34E-04 0.692 -.116 0.116                            0.036  127    38
 2  2.5  5.0   0.0 1.07E-04 0.698 0.259 0.074                            0.036  127    36
 3  2.5  2.5   0.0 3.58E-05 0.744 0.308 0.056                            0.036  376    35
 4  1.2  2.5   0.0 7.99E-05 0.731 -.043 0.089                            0.036  376    38
 5  1.2  5.0  -2.5 2.40E-04 1.419 -.249 0.134                            0.036  127    32
 6  2.5  5.0  -2.5 1.05E-04 1.494 0.037 0.078                            0.036  127    29
 7  2.5  2.5  -2.5 3.45E-05 1.588 0.079 0.062                            0.036  376    28
 8  1.2  2.5  -2.5 7.98E-05 1.514 -.231 0.105                            0.036  376    30
/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\
SAT#1 GIVEN VTL, DELTAL, THETAL, & RL, FIND BETAS, EPSILONS, & LAMBDAS
K%   L    W    VB    BETA     VT   DELTA                   EPSILON LAMBDA THETA   R    DATA
 1  1.2  5.0   0.0 2.64E-04 0.692 -.116                     0.620  0.047 0.036  127    36
 2  2.5  5.0   0.0 1.15E-04 0.698 0.259                     0.174  0.023 0.036  127    38
 3  2.5  2.5   0.0 3.89E-05 0.744 0.308                     0.129  0.021 0.036  376    39
 4  1.2  2.5   0.0 8.94E-05 0.731 -.043                     0.500  0.041 0.036  376    36
 5  1.2  5.0  -2.5 2.82E-04 1.419 -.249                     0.712  0.056 0.036  127    34
 6  2.5  5.0  -2.5 1.12E-04 1.494 0.037                     0.189  0.022 0.036  127    37
 7  2.5  2.5  -2.5 3.89E-05 1.588 0.079                     0.172  0.016 0.036  376    30
 8  1.2  2.5  -2.5 9.00E-05 1.514 -.231                     0.558  0.050 0.036  376    28
/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\
SAT#2 GIVEN VTL, BETAS0, DELTAL, LAMBDAS, THETAL, & RL, FIND EPSILONSB
K%   L    W    VB    BETA     VT   DELTA                   EPSILON LAMBDA THETA   R    DATA
 5  1.2  5.0  -2.5 2.64E-04 1.419 -.249                     0.634  0.056 0.036  127    34
 6  2.5  5.0  -2.5 1.15E-04 1.494 0.037                     0.210  0.022 0.036  127    37
 7  2.5  2.5  -2.5 3.89E-05 1.588 0.079                     0.171  0.016 0.036  376    30
 8  1.2  2.5  -2.5 8.94E-05 1.514 -.231                     0.550  0.050 0.036  376    28
/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\
LIN#2 GIVEN VTL, BETAS0, DELTAL, EPSILONS0&SB, THETAL, & RL, FIND ETAL
K%   L    W    VB    BETA     VT   DELTA             ETA EPSILON LAMBDA THETA   R    DATA
 1  1.2  5.0   0.0 2.64E-04 0.692 -.116            1.007 0.620  0.047 0.036  127    38
 2  2.5  5.0   0.0 1.15E-04 0.698 0.259            0.531 0.174  0.023 0.036  127    36
 3  2.5  2.5   0.0 3.89E-05 0.744 0.308            0.305 0.129  0.021 0.036  376    35
 4  1.2  2.5   0.0 8.94E-05 0.731 -.043            0.532 0.500  0.041 0.036  376    38
 5  1.2  5.0  -2.5 2.64E-04 1.419 -.249            1.081 0.634  0.056 0.036  127    32
 6  2.5  5.0  -2.5 1.15E-04 1.494 0.037            0.653 0.210  0.022 0.036  127    29
 7  2.5  2.5  -2.5 3.89E-05 1.588 0.079            0.526 0.171  0.016 0.036  376    28
 8  1.2  2.5  -2.5 8.94E-05 1.514 -.231            0.723 0.550  0.050 0.036  376    30
/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\/\
```

Table 2.3.6-4. Test Results for 1.2-μm n-MOSFET (CMOS1N1)

INDIVIDUAL XT FIT ANALYSIS

K%	L	W	VB	BETA	VT	DELTA	TAU	ETA	EPSILON	LAMBDA	THETA	R	CC
1	1.2	5.0	0.0	2.64E-04	0.692	-.116	0.116	1.007	0.620	0.047	0.036	127	.9994
2	2.5	5.0	0.0	1.15E-04	0.698	0.259	0.074	0.531	0.174	0.023	0.036	127	.9998
3	2.5	2.5	0.0	3.89E-05	0.744	0.308	0.056	0.305	0.129	0.021	0.036	376	.9999
4	1.2	2.5	0.0	8.94E-05	0.731	-.043	0.089	0.532	0.500	0.041	0.036	376	.9999
5	1.2	5.0	-2.5	2.64E-04	1.419	-.249	0.134	1.081	0.634	0.056	0.036	127	.9993
6	2.5	5.0	-2.5	1.15E-04	1.494	0.037	0.078	0.653	0.210	0.022	0.036	127	.9997
7	2.5	2.5	-2.5	3.89E-05	1.588	0.079	0.062	0.526	0.171	0.016	0.036	376	.9997
8	1.2	2.5	-2.5	8.94E-05	1.514	-.231	0.105	0.723	0.550	0.050	0.036	376	.9996

SUMMARY OF GLOBAL TRANSISTOR PARAMETERS.

BETA:	KP(A/V^2)	= .707611E-04
	DELTA W(um)	= 1.22189
	DELTA L(um)	= .180254
THRESHOLD:	VT0(V)	= .715458
	PSI(V)	= .100833
	GAMMA(V^0.5)	= .793478
	KLG(um)	= .431201E-01
	KWG(um)	= .617277E-01
DELTA:	D0(V^0.5)	= .31999
	KLD(um)	= .532337
	KWD(um)	= .146257
TAU:	THETA(1/V)	= .360062E-01
	KLT(um/V)	= .680515E-01
	RW(Ohm*um)	= 480.854
ETA:	H0(1/V)	= .243186
	KLH(um/V)	= .604351
EPSILON:	E0(1/V)	= .14657
	KLE(um/V)	= .736832
LAMBDA:	L0(1/V)	= .28799E-02
	KLL(um/V)	= .520684E-01
	KWL(um/V)	= .997387E-02

GLOBAL XT FIT ANALYSIS

K%	L	W	VB	BETA	VT	DELTA	TAU	ETA	EPSILON	LAMBDA	THETA	R	CC
1	1.2	5.0	0.0	2.62E-04	0.700	-.070	0.103	0.836	0.576	0.051	0.036	127	.9997
2	2.5	5.0	0.0	1.15E-04	0.714	0.222	0.065	0.504	0.171	0.023	0.036	127	.9997
3	2.5	2.5	0.0	3.90E-05	0.733	0.298	0.065	0.504	0.171	0.018	0.036	376	.9973
4	1.2	2.5	0.0	8.87E-05	0.719	0.006	0.103	0.836	0.576	0.046	0.036	376	.9917
5	1.2	5.0	-2.5	2.62E-04	1.417	-.302	0.103	0.836	0.576	0.051	0.036	127	.9960
6	2.5	5.0	-2.5	1.15E-04	1.491	-.009	0.065	0.504	0.171	0.023	0.036	127	.9938
7	2.5	2.5	-2.5	3.90E-05	1.590	0.067	0.065	0.504	0.171	0.018	0.036	376	.9997
8	1.2	2.5	-2.5	8.87E-05	1.517	-.226	0.103	0.836	0.576	0.046	0.036	376	.9977

Table 2.3.6-5. Test Results for 1.2-μm p-MOSFET (CM01P4S)

INDIVIDUAL XT FIT ANALYSIS

K%	L	W	VB	BETA	VT	DELTA	TAU	ETA	EPSILON	LAMBDA	THETA	R	CC
1	1.2	5.0	0.0	8.75E-05	1.120	0.078	0.136	0.658	0.164	0.109	0.097	146	.9997
2	2.5	5.0	0.0	3.44E-05	1.108	0.277	0.119	0.383	0.027	0.047	0.097	146	.9998
3	2.5	2.5	0.0	1.33E-05	1.067	0.345	0.096	0.329	0.009	0.044	0.097	404	.9998
4	1.2	2.5	0.0	2.97E-05	1.067	0.169	0.109	0.510	0.117	0.098	0.097	404	.9997
5	1.2	5.0	-2.5	8.75E-05	1.659	-.083	0.137	0.894	0.206	0.127	0.097	146	.9995
6	2.5	5.0	-2.5	3.44E-05	1.760	0.104	0.103	0.818	0.087	0.045	0.097	146	.9994
7	2.5	2.5	-2.5	1.33E-05	1.746	0.160	0.069	0.958	0.091	0.044	0.097	404	.9990
8	1.2	2.5	-2.5	2.97E-05	1.638	0.013	0.098	0.924	0.169	0.111	0.097	404	.9994

SUMMARY OF GLOBAL TRANSISTOR PARAMETERS.

BETA:	KP(A/V^2)	= .191164E-04
	DELTA W(um)	= 1.08108
	DELTA L(um)	= .337993
THRESHOLD:	VT0(V)	= 1.11823
	PSI(V)	= .548406
	GAMMA(V^0.5)	= .735643
	KLG(um)	= .046075
	KWG(um)	=-.181214E-01
DELTA:	D0(V^0.5)	= .249434
	KLD(um)	= .239326
	KWD(um)	= .194178
TAU:	THETA(1/V)	= .973776E-01
	KLT(um/V)	= .219421E-01
	RW(Ohm*um)	= 573.908
ETA:	H0(1/V)	= .539402
	KLH(um/V)	= .178574
EPSILON:	E0(1/V)	= .197648E-01
	KLE(um/V)	= .158625
LAMBDA:	L0(1/V)	= .947916E-02
	KLL(um/V)	= .951203E-01
	KWL(um/V)	= .177217E-01

GLOBAL XT FIT ANALYSIS

K%	L	W	VB	BETA	VT	DELTA	TAU	ETA	EPSILON	LAMBDA	THETA	R	CC
1	1.2	5.0	0.0	8.69E-05	1.083	0.094	0.123	0.747	0.164	0.115	0.097	146	.9994
2	2.5	5.0	0.0	3.47E-05	1.103	0.261	0.108	0.622	0.054	0.049	0.097	146	.9990
3	2.5	2.5	0.0	1.25E-05	1.098	0.348	0.108	0.622	0.054	0.041	0.097	404	.9821
4	1.2	2.5	0.0	3.15E-05	1.078	0.181	0.123	0.747	0.164	0.107	0.097	404	.9995
5	1.2	5.0	-2.5	8.69E-05	1.664	-.086	0.123	0.747	0.164	0.115	0.097	146	.9983
6	2.5	5.0	-2.5	3.47E-05	1.763	0.081	0.108	0.622	0.054	0.049	0.097	146	.9941
7	2.5	2.5	-2.5	1.25E-05	1.738	0.168	0.108	0.622	0.054	0.041	0.097	404	.9994
8	1.2	2.5	-2.5	3.15E-05	1.638	0.001	0.123	0.747	0.164	0.107	0.097	404	.9943

Table 2.3.6-6. Test Results for 3-μm n-MOSFET (RADN.DAT)

INDIVIDUAL XT FIT ANALYSIS

K%	L	W	VB	BETA	VT	DELTA	TAU	ETA	EPSILON	LAMBDA	THETA	R	CC
1	3.0	9.0	0.0	1.70E-04	0.651	0.273	0.061	0.525	0.172	0.035	0.038	47	.9998
2	9.0	9.0	0.0	4.91E-05	0.661	0.452	0.047	0.255	0.000	0.013	0.038	47	.9999
3	9.0	4.5	0.0	2.03E-05	0.671	0.482	0.038	0.175	-.008	0.014	0.038	114	.9999
4	3.0	4.5	0.0	6.94E-05	0.675	0.295	0.047	0.338	0.147	0.032	0.038	114	.9999
5	3.0	9.0	-2.5	1.70E-04	1.417	0.018	0.055	0.625	0.214	0.038	0.038	47	.9996
6	9.0	9.0	-2.5	4.91E-05	1.488	0.174	0.041	0.398	0.046	0.013	0.038	47	.9998
7	9.0	4.5	-2.5	2.03E-05	1.520	0.199	0.030	0.405	0.046	0.012	0.038	114	.9997
8	3.0	4.5	-2.5	6.94E-05	1.486	0.033	0.043	0.518	0.199	0.035	0.038	114	.9997

SUMMARY OF GLOBAL TRANSISTOR PARAMETERS.

BETA:	KP(A/V^2)	= .543281E-04
	DELTA W(um)	= 1.36203
	DELTA L(um)	= .556277
THRESHOLD:	VT0(V)	= .661496
	PSI(V)	= .324927E-01
	GAMMA(V^0.5)	= .812039
	KLG(um)	= .567158E-01
	KWG(um)	= .889447E-01
DELTA:	D0(V^0.5)	= .382444
	KLD(um)	= .574019
	KWD(um)	= .157049
TAU:	THETA(1/V)	= .381296E-01
	KLT(um/V)	= .390104E-01
	RW(Ohm*um)	= 359.026
ETA:	H0(1/V)	= .229779
	KLH(um/V)	= .663555
EPSILON:	E0(1/V)	= .450726E-01
	KLE(um/V)	= .557218
LAMBDA:	L0(1/V)	= .577087E-02
	KLL(um/V)	= .074795
	KWL(um/V)	= .743894E-02

GLOBAL XT FIT ANALYSIS

K%	L	W	VB	BETA	VT	DELTA	TAU	ETA	EPSILON	LAMBDA	THETA	R	CC
1	3.0	9.0	0.0	1.70E-04	0.655	0.279	0.054	0.501	0.183	0.035	0.038	47	.9994
2	9.0	9.0	0.0	4.91E-05	0.664	0.446	0.043	0.308	0.021	0.014	0.038	47	.9988
3	9.0	4.5	0.0	2.02E-05	0.674	0.476	0.043	0.308	0.021	0.012	0.038	114	.9966
4	3.0	4.5	0.0	6.98E-05	0.665	0.309	0.054	0.501	0.183	0.034	0.038	114	.9975
5	3.0	9.0	-2.5	1.70E-04	1.426	0.003	0.054	0.501	0.183	0.035	0.038	47	.9981
6	9.0	9.0	-2.5	4.91E-05	1.478	0.170	0.043	0.308	0.021	0.014	0.038	47	.9967
7	9.0	4.5	-2.5	2.02E-05	1.529	0.199	0.043	0.308	0.021	0.012	0.038	114	.9985
8	3.0	4.5	-2.5	6.98E-05	1.478	0.032	0.054	0.501	0.183	0.034	0.038	114	.9989

Table 2.3.6-7. Test Results for 3-μm p-MOSFET (RADP.DAT)

INDIVIDUAL XT FIT ANALYSIS

K%	L	W	VB	BETA	VT	DELTA	TAU	ETA	EPSILON	LAMBDA	THETA	R	CC
1	3.0	9.0	0.0	5.94E-05	0.735	0.208	0.126	0.525	0.043	0.062	0.096	167	.9998
2	9.0	9.0	0.0	1.77E-05	0.778	0.299	0.110	0.294	-.020	0.024	0.096	167	.9998
3	9.0	4.5	0.0	7.31E-06	0.842	0.323	0.093	0.143	-.033	0.023	0.096	407	.9999
4	3.0	4.5	0.0	2.43E-05	0.815	0.228	0.105	0.354	0.026	0.057	0.096	407	.9998
5	3.0	9.0	-2.5	5.94E-05	1.232	0.050	0.111	0.905	0.104	0.061	0.096	167	.9993
6	9.0	9.0	-2.5	1.77E-05	1.310	0.123	0.092	0.704	0.051	0.021	0.096	167	.9994
7	9.0	4.5	-2.5	7.31E-06	1.391	0.156	0.065	0.670	0.046	0.019	0.096	407	.9992
8	3.0	4.5	-2.5	2.43E-05	1.332	0.062	0.083	0.823	0.101	0.056	0.096	407	.9992

SUMMARY OF GLOBAL TRANSISTOR PARAMETERS.

BETA:	KP(A/V^2)	= .198667E-04
	DELTA W(um)	= 1.35423
	DELTA L(um)	= .442207
THRESHOLD:	VT0(V)	= .782031
	PSI(V)	= .405029
	GAMMA(V^0.5)	= .486708
	KLG(um)	= .854485E-01
	KWG(um)	= .174861
DELTA:	D0(V^0.5)	= .239461
	KLD(um)	= .303391
	KWD(um)	= .15217
TAU:	THETA(1/V)	= .956032E-01
	KLT(um/V)	= .508118E-01
	RW(Ohm*um)	= 1278.82
ETA:	H0(1/V)	= .367785
	KLH(um/V)	= .726861
EPSILON:	E0(1/V)	= .133359E-01
	KLE(um/V)	= .209502
LAMBDA:	L0(1/V)	= .960858E-02
	KLL(um/V)	= .136847
	KWL(um/V)	= .180511E-01

GLOBAL XT FIT ANALYSIS

K%	L	W	VB	BETA	VT	DELTA	TAU	ETA	EPSILON	LAMBDA	THETA	R	CC
1	3.0	9.0	0.0	5.94E-05	0.776	0.210	0.115	0.652	0.069	0.061	0.096	167	.9960
2	9.0	9.0	0.0	1.77E-05	0.790	0.294	0.102	0.453	0.011	0.023	0.096	167	.9970
3	9.0	4.5	0.0	7.30E-06	0.809	0.322	0.102	0.453	0.011	0.020	0.096	407	.9960
4	3.0	4.5	0.0	2.44E-05	0.795	0.239	0.115	0.652	0.069	0.057	0.096	407	.9964
5	3.0	9.0	-2.5	5.94E-05	1.229	0.037	0.115	0.652	0.069	0.061	0.096	167	.9953
6	9.0	9.0	-2.5	1.77E-05	1.302	0.120	0.102	0.453	0.011	0.023	0.096	167	.9930
7	9.0	4.5	-2.5	7.30E-06	1.403	0.149	0.102	0.453	0.011	0.020	0.096	407	.9960
8	3.0	4.5	-2.5	2.44E-05	1.331	0.066	0.115	0.652	0.069	0.057	0.096	407	.9962

Table 2.3.6-8. Test Results for n-MOSFETs (FIT81NP.BAS)

TECHNOLOGY (μm)	1.2	1.2	3.0	3.0
MANUFACTURER	A	A	B	C
NO. TRANSISTORS	4	8	4	4
|VB(V)|	0, 2.5	0, 2.5, 5	0, 2.5	0, 2.5
FILE	CMOS1N1	CMOS1N4	RADN	KELN04
KP ($\mu A/V^2$)	70.7	64.1	54.3	48.8
DELTA W (μm)	1.222	1.166	1.362	0.692
DELTA L (μm)	0.180	0.141	0.556	1.110
VTO (V)	0.715	0.844	0.661	0.858
PSI (V)	0.101	0.160	0.032	-0.113
GAMMA ($V^{0.5}$)	0.794	0.883	0.812	1.254
KLG (μm)	0.043	0.017	0.057	0.002
KWG (μm)	0.062	0.041	0.089	-0.114
DO ($V^{0.5}$)	0.320	0.304	0.382	0.482
KLD (μm)	0.532	0.459	0.574	0.593
KWD (μm)	0.146	0.200	0.157	0.474
THETA (1/V)	0.036	0.033	0.038	0.042
KLT (μm/V)	0.068	0.078	0.039	0.045
RW (ohm*μm)	480.0	609.0	359.0	457.0
HO (1/V)	0.243	0.611	0.230	0.346
KLH (μm/V)	0.605	0.170	0.664	0.406
EO (1/V)	0.147	0.059	0.045	0.057
KLE (μm/V)	0.737	0.600	0.557	0.551
LO (1/V)	0.003	0.001	0.006	0.002
KLL (μm/V)	0.052	0.062	0.075	0.078
KWL (μm/V)	0.010	0.009	0.007	0.003

Table 2.3.6-9. Test Results for p-MOSFETs (FIT81NP.BAS)

TECHNOLOGY (μm)	1.2	1.2	3.0
MANUFACTURER	A	A	B
NO. TRANSISTORS	4	8	4
|VB(V)|	0, 2.5	0, 2.5, 5	0, 2.5
FILE	CMOS1P4S	CMOS1N4	RADP
KP ($\mu A/V^2$)	19.1	21.1	19.9
DELTA W (μm)	1.081	1.218	1.354
DELTA L (μm)	0.338	0.283	0.442
VTO (V)	1.118	1.116	0.782
PSI (V)	0.548	0.543	0.405
GAMMA ($V^{0.5}$)	0.736	0.740	0.487
KLG (μm)	0.046	0.062	0.085
KWG (μm)	-0.018	0.008	0.175
DO ($V^{0.5}$)	0.249	0.273	0.239
KLD (μm)	0.239	0.261	0.303
KWD (μm)	0.194	0.155	0.152
THETA (1/V)	0.097	0.100	0.096
KLT (μm/V)	0.022	0.026	0.051
RW (ohm*μm)	573.0	620.0	1279.0
HO (1/V)	0.539	0.750	0.368
KLH (μm/V)	0.179	0.076	0.727
EO (1/V)	0.020	-0.007	0.013
KLE (μm/V)	0.159	0.163	0.210
LO (1/V)	0.009	0.002	0.010
KLL (μm/V)	0.095	0.112	0.137
KWL (μm/V)	0.018	0.014	0.018

2.4 TEST CHIP METHODOLOGY

In this section the various test chips developed on this program are mentioned, and two of the 3-μm CMOS bulk test chips are evaluated with respect to the area required by the test structures and the test time to evaluate the chips. It was found that a reasonably comprehensive test chip takes an area of 6.0 mm x 6.4 mm and requires 10 minutes to test. In the latter part of this section, our test chip assembler and test program generator are discussed, along with our statistical analysis package.

2.4.1 1.2-μm CMOS Technology[1]

2.4.1.1 Test Structure Designs. A set of test structures was designed using the JPL test chip assembler and was used to evaluate the first CMOS-bulk foundry runs with feature sizes of 1.2 μm (for comparison purposes note that standard CMOS-bulk fabrication runs use a minimum feature size of 3 μm). In addition to the problems associated with the physical scaling of the structures, this geometry provided an additional set of unique problems, since the design files had to be generated in such a way as to be capable of being processed through p-well, n-well, and twin-well processing lines. This requirement meant that the files containing the geometrical design rules as well as the structure design files had to produce process-insensitive designs, a requirement that does not apply to the more mature 3.0-μm CMOS feature size technology. Because of the photolithographic steps required with this feature size, the maximum allowable chip size was 10 mm x 10 mm, and this chip was divided into 24 project areas, with each area being 1.6 mm x 1.6 mm in size. The JPL-designed structures occupied 13 out of the 21 allowable project sizes and provided the only test information obtained from these three preliminary runs. The structures were used to successfully evaluate three different manufacturing runs through two separate foundries.

2.4.1.2 Test Structure Evaluations. The JPL LSI-11/23 test system was used to conduct the preliminary wafer evaluations of the fabrication runs. The first wafers were received on April 29, 1984, and the test information and summary data were supplied to the Information Sciences Institute (ISI) approximately ten days later. For the complete wafer evaluations, test programs were written for the ISI Accutest test system. This software was used to successfully evaluate the remainder of the runs and is now in place at ISI for subsequent evaluations.

2.4.1.3 Data Analysis. The data obtained from these runs was analyzed using a statistical package which is described in Section 2.4.5.

2.4.1.4 Results and Conclusions. The results of the 1.2-μm CMOS foundry runs indicate that a comparatively simple set of design rules can be used to fabricate structures and circuits within the constraints of a foundry environment as used for the more mature 3-μm CMOS and NMOS technologies. The test data also indicate that at the present time this technology is approximately 1-2 years from being mature enough for routine foundry fabrication runs. This is due to the lack of a uniform, scalable set of design rules that can be used through several manufacturers, the inability to use multi-project chips (routinely processed through the 3-μm CMOS-bulk lines), and the lack of adequate process controls for some parameters. Despite these shortcomings, the technology is moving along quite rapidly and there appear to be no major technical obstacles to be overcome.

[1]This section was prepared by C. A. Pina.

2.4.2 3-μm CMOS Technology[1]

During this period, a number of test chips were designed and submitted for fabrication. A list of these test chips and their purpose is provided in Table 2.4.2-1. The primary purpose of these designs was to obtain the structures and the methods needed to gather process parametric and yield information. The chips also included experimental structures and structures needed to obtain data for fault modeling (faulted NAND gates, transistors, and inverters) structures permitting hot-electron injection.

The result of the design of these test chips has been the development of a set of test chips which are discussed in Section 2.4.3. These test chips have been developed for use in lot acceptance of LSI/VLSI integrated circuits and the qualification of wafer fabrication processes.

[1]This section was prepared by C. A. Pina.

Table 2.4.2-1. JPL Test Chip Summary (October 10, 1984)[a]

Test Chip	Description
CM4094	Parametric (small)
CM4092	Parametric/yield
CM4062	Parametric/yield
CM4071A	Parametric (CMOS125)
CM4071C	Parametric (CMOS125)
CM4071D	Parametric (CMOS125)
CM4071T	Parametric (CMOS125)
CM4061	Parametric/yield
CM4042	Parametric/yield
CM4031A	Reliability
CM4031B	Reliability
CM4031C	Reliability
CM4013A	Reliability
CM4014A	Reliability
CM4012	Parametric/yield
CM4013	Reliability
CM4014	Reliability
CRESCHP1	CRRES prototype
CRESCHP1A	CRRES prototype
CRESCHP2	CRRES prototype
CRRESCHP7	CRRES prototype
CRRESCHP8	CRRES prototype
CRRESCHP10	CRRES prototype
CRRESCHP11	CRRES prototype

[a]Note that these chips were fabricated in 3-μm CMOS-bulk (p-well), except for CM4071A, CM4071C, CM4071D, and CM4071T.

2.4.3 3-μm CMOS-Bulk Test Chips

In this section, we characterize two 3-μm CMOS bulk test chips in terms of the area devoted to each of the test structures and the test time required to evaluate the chips. The area and test times were taken from the data presented in Section 2.2.2 for each of the test structures. The chips CM5031 and CM5041 are shown in Figures 2.4.3-1 and -2. These chips were not designed to include all the possible test structures. Instead, they were designed to provide a major portion of the information needed for wafer acceptance.

These chips have a similar set of test structures. The major difference between the two chips may be seen in the presence of the TDDB structure that appears in the CM5041 test chip. The complement of test structures found in each of the chips is indicated in Table 2.4.3-1, where it is seen that only certain parameters are measured. The emphasis has been to measure those parameters that characterize the implant and metal layers, the DC device parameters, the DC and AC circuit parameters, the layer line-width and spacing, the defect densities, and some reliability parameters.

The number of times that each test structure is placed in the test chip depends on a variety of factors. For the SXBR, the number depends on measuring each of the conducting layers in a CMOS process, i.e., p-diffusion, n-diffusion, p-poly on thin oxide, n-poly on thin oxide, and metal. In addition, the structures have been repeated with different line-widths. For the CR, the number reflects the evaluation of both p-poly-metal, n-poly-metal, p-diffusion-metal and n-diffusion-metal contacts of various sizes including undersize contacts to determine the contact-process cliff. For the TR, the number reflects the evaluation of transistors of various channel widths and lengths as might be encountered in circuit design.

The area required by the various test structures is listed in Table 2.4.3-2. The test structures are listed in the most appropriate category in the table even though they could be listed under a number of categories. For instance, the SXBR is listed in the Process Parameters even though it could also be listed under the Device Parameters and the Layout Rule Parameters. This prevents double counting the test structure area and test time. Of course, the structures that consume the largest amount of area are the INV-A, STP/CMB, CAP-A, and TDDB, although a significant amount of area is also devoted to the SXBR, CR, RTTR, and TRs. Notice that each of the latter set of test structures is designed to be a set of stand-alone structures, which means that no common busses are shared between the structures. This is the best approach to test structure layout for wafer acceptance, for the stand-alone approach, as opposed to the common bussing approach, eliminates the possibility that the evaluation of one structure will be affected by another structure.

The overall area for the test structures is given at the bottom of Table 2.4.3-2, along with the overall chip size. Notice that the area consumed by the test structures is less than the chip area. This is understandable since the area used for each test structure does not account for the space

[1]This section was prepared by C. A. Pina.

between the probe pads. In addition, some portions of the test chip has not been fully utilized. Once the chips have been tested and proven, a more compact version of the chips will be designed.

The test time is also listed in Table 2.4.3-2. As explained in Section 2.2.2, these test times are the minimum test times to evaluate each test structure. The table indicates that the longest test times occur for the INV-A and the TDDB test structures, although the the 6TINV and the TRs also take a significant amount of test time. The overall test time for these chips is substantial. As indicated at the bottom of Table 2.4.3-2, it is 10 to 20 minutes per chip. Note that these test times do not include the time for prober motion; however, this time is a small fraction of the overall test time.

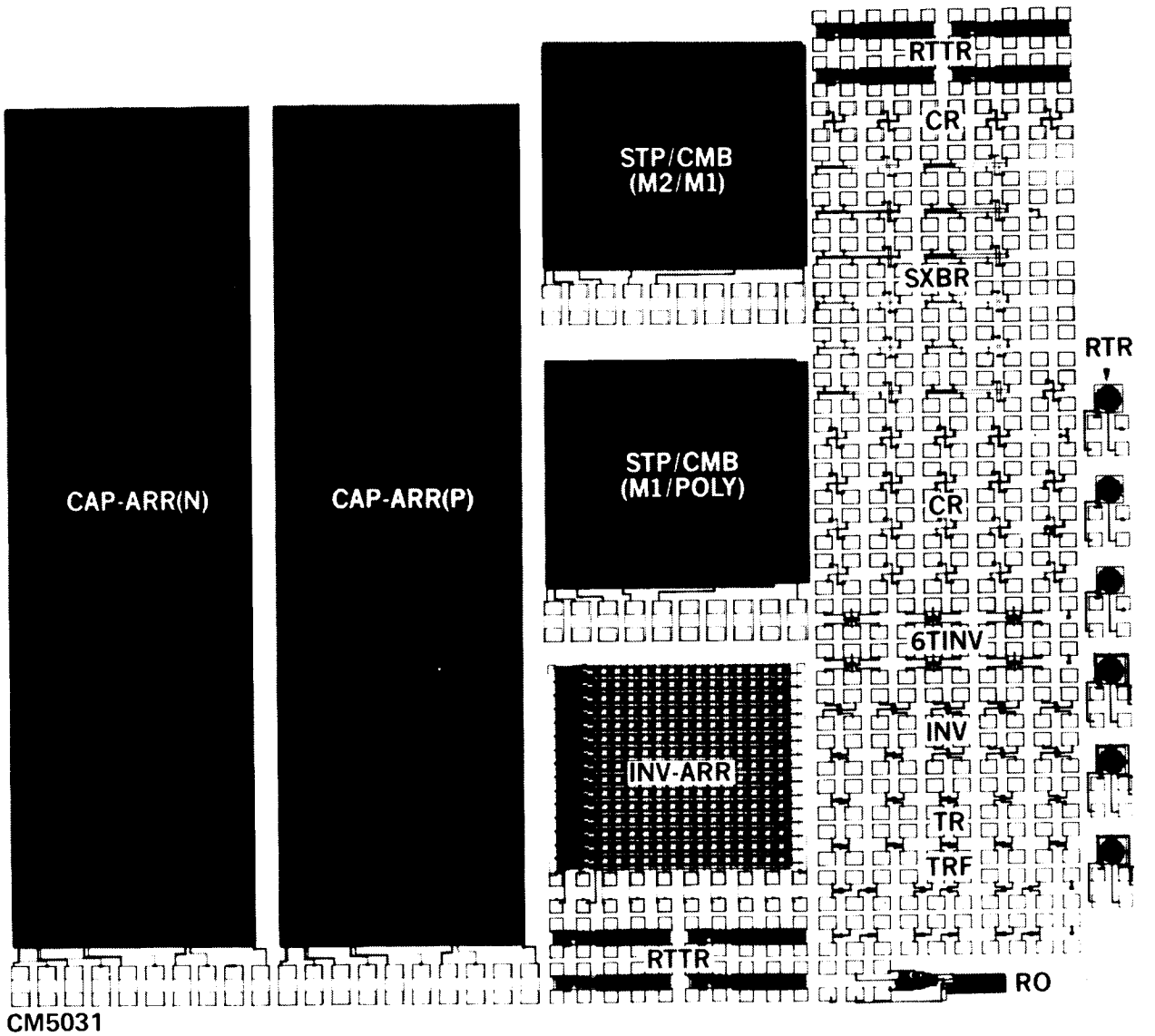


Figure 2.4.3-1. Test chip CM5031 (area = 6.0 mm x 6.4 mm)

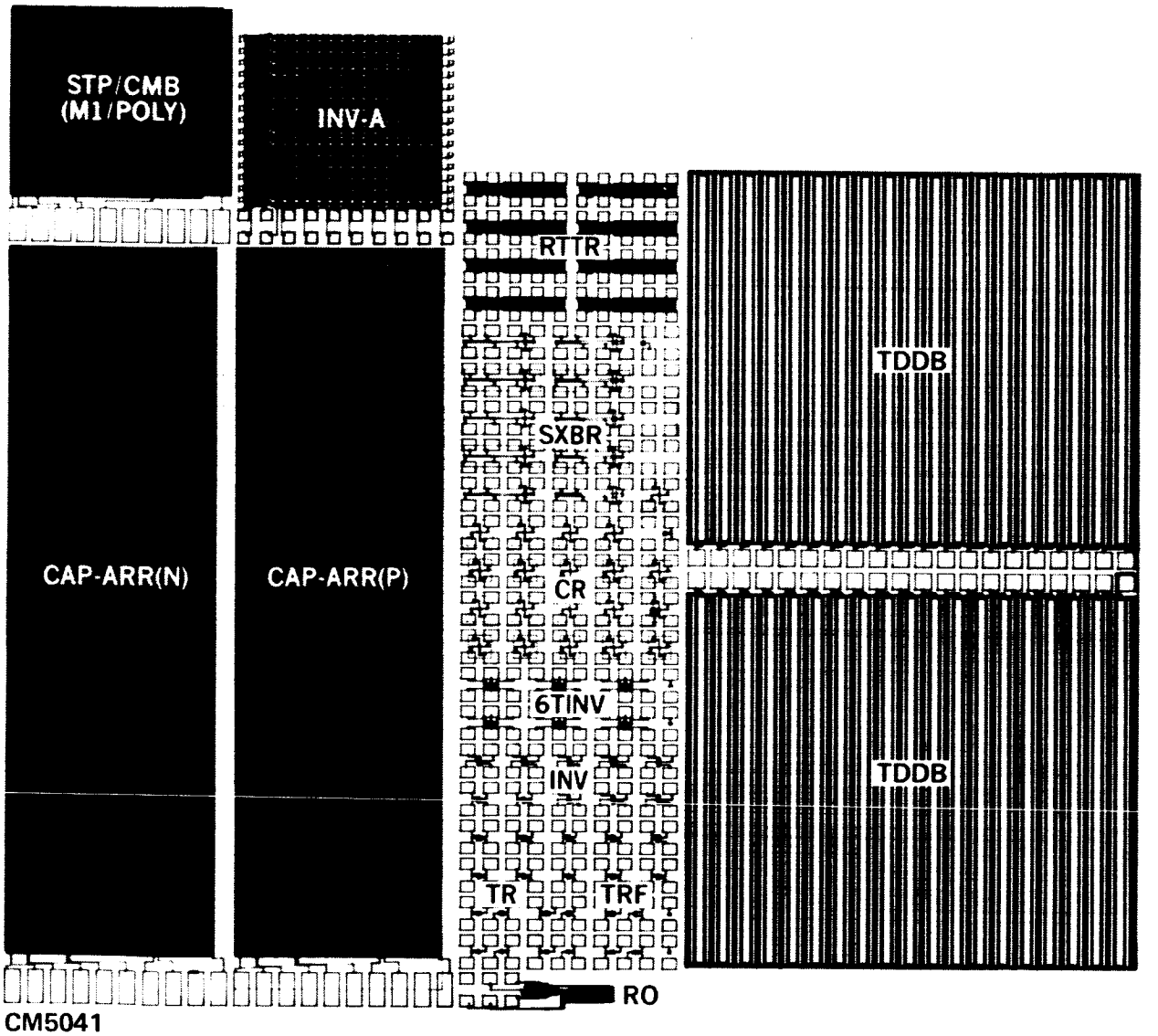


Figure 2.4.3-2. Test chip CM5041 (area = 7.2 mm x 8.0 mm)

Table 2.4.3-1. Chip Coverage of the Critical Parameter Set

TEST STRUCTURES PARAMETERS	TEST STRUCTURES	CHIP 5031		CHIP 5041	
1.0 PROCESS PARAMETERS					
1.1 Layer Line-Width	XBR, SXBR		X		X
1.2 Layer Sheet Resistance	XBR, SXBR		X		X
1.3 Metal-Layer Contact Resistance	CR, CR-A	X		X	
1.4 Body Dopant Density	CAP, TR				
1.5 Body Resistivity	PFPR				
1.6 Layer-Layer Alignment	ALIR				
1.7 Body Lifetime	DI, CAP				
1.8 Junction Leakage Current (IDBLEAK)	TR	X		X	
1.9 Junction Breakdown Voltage	TR, DI	X		X	
1.10 Oxide Thickness	CAP, CAPF, RTTF		X		X
1.11 Gate-Oxide Breakdown Voltage (VBG)	CAP, RTTR		X		X
1.12 Field-Oxide Threshold Voltage	TRF	X	X		
1.13 Channel Leakage Current (IDSO)	TR	X	X		
1.14 Transistor Punch-Through Voltage (VPT)	TR	X	X		
2.0 DEVICE PARAMETERS					
2.1 TRANSISTORS					
2.1.1 Threshold Voltage (VTO)	TR, 6TINV	X	X	X	X
2.1.2 Body Effect Factor (GAMMA)	TR, 6TINV	X	X	X	X
2.1.3 Conduction Factor (KP)	TR, 6TINV	X	X	X	X
2.1.4 Effective Channel Width (WE)	TR, 6TINV	X	X	X	X
2.1.5 Effective Channel Length (LE)	TR, 6TINV	X	X	X	X
2.1.6 Channel Length Modulation (LAMBDA)	TR, 6TINV	X	X	X	X
2.1.7 Subthreshold Conduction Factor (NFS)	TR, 6TINV	X	X	X	X
2.1.8 Gate Oxide Capacitance (COX)	RTTR, CAP	X		X	
2.1.9 Gate-Source/Drain Overlap Cap. (CG/DSO)	RTTR	X		X	
2.1.10 Gate-Body Overlap Capacitance (CGBO)	CAPF				
2.1.11 Zero-Bias Junction Capacitance (CJ)	RTDI				
2.1.12 Zero-Bias Junction Grading Factor (MJ)	RTDI				
2.1.13 Zero-Bias Sidewall Junction Cap. (CJSW)	RTDI				
2.1.14 Zero-B. Sw. Junc. Cap. Grad. Fact. (MJSW)	RTDI				
2.2 CONTACTS					
2.2.1 Contact Resistance	CR	X		X	
2.3 WIRES					
2.3.1 Layer Sheet Resistance	XBR, SXBR		X		X
2.3.2 Layer Line-Width	XBR, SXBR		X		X
2.3.3 Layer-Layer Capacitance	STP, CMB		X		X
3.0 CIRCUIT PARAMETERS					
3.1 Inverter Output for Input Low (VHIGH)	INV, INV-A	X	X	X	X
3.2 Inverter Output for Input High (VLOW)	INV, INV-A	X	X	X	X
3.3 Inverter Threshold for Output = Input (VINV)	INV, INV-A	X	X	X	X
3.4 Inverter Gain (GAIN)	INV, INV-A	X	X	X	X
3.5 Inverter Noise Margin (INVNM)	INV, INV-A	X	X	X	X
3.6 Gate Delay (TAU)	RO, TS	X		X	

Table 2.4.3-1. (continued)

TEST STRUCTURES PARAMETERS	TEST STRUCTURES	CHIP 5031	CHIP 5041
4.0 LAYOUT RULE PARAMETERS			
4.1 Layer Line-Width	XBR, SXBR	X	X
4.2 Same Layer Spacing	SXBR, CS	X	X
4.3 Different Layer Spacing	CS, ALIR		
4.4 Contact Size	CR, CR-A		
4.5 Layer Extensions	CS		
5.0 DEFECT PARAMETERS			
5.1 DEFECT DENSITY			
5.1.1 Shorts: Different Layers (Pinholes)	CAP-A	X	X
5.1.2 Opens: Different Layers (Contact Integrity)	CR, CR-A	X	X
5.1.3 Shorts: Same Layer (Layer Bridging)	CMB	X	X
5.1.4 Opens: Same Layer (Layer Step Coverage)	STP	X	X
5.2 DEFECT LOCATION			
5.2.1 Transistor Matrix	TR-A		
5.2.2 Inverter Matrix	INV-A		
5.2.3 Static RAM	SRAM		
6.0 RELIABILITY PARAMETERS			
6.1 Time-Dependent Dielectric Breakdown	TDDB	X	
6.2 Radiation Hardness	RTR, TR	X	X
6.3 Electromigration	CR, CMB	X	X
6.4 Oxide Instabilities	TR, CAP	X	X
6.5 Contact Integrity	CR, CR-A	X	X
6.6 Latch-Up Transistor	LUTR		

Table 2.4.3-2. Test Chip Area and Test Time

TEST STRUCTURE	CHIP CM5031 NUMBER/AREA/TEST TIME (#/mm^2/seconds)			CHIP CM5041 NUMBER/AREA/TEST TIME (#/mm^2/seconds)		
PROCESS PARAMETERS						
SXBR	12	1.61	12.0	10	1.34	10.0
CR	25	1.44	12.5	20	1.15	10.0
TRF	2	0.16	1.0	2	0.16	1.0
DEVICE PARAMETERS						
RTTR	6	1.07	6.0	6	1.07	6.0
6TINV	6	0.58	60.0	6	0.58	60.0
TR	20	1.15	100.0	20	1.15	100.0
CIRCUIT PARAMETERS						
INV	9	0.52	9.0	9	0.52	9.0
INV-A	1	2.31	300.0	1	2.31	300.0
RO	1	0.27	0.5	1	0.27	0.5
LAYOUT RULE PARAMETERS						
DEFECT DENSITY PARAMETERS						
STP/CMB	2	4.62	2.0	1	2.31	2.0
CAP-A	2	20.79	10.0	2	16.17	10.0
RELIABILITY PARAMETERS						
RTR	6	0.50	90.0			
TDDB				1	18.48	600.0
STRUCTURE TOTALS	92	34.30	603.0	79	45.24	1108.5
TOTAL CHIP AREA		38.4			57.6	

2.4.4 Test Chip Assembler and Test Program Generator[1]

One of the major problems in working at the geometry level for the generation of either test structure or functional circuit designs is the amount of labor involved in the design phase. In the case of test structure design, the designer has to stay reasonably close to the geometry level, since many of the parameters he is trying to evaluate are intimately related to the geometrical design rules used in the particular fabrication process being used. Similarly, the generation of test programs to evaluate the structures is a labor-intensive process if such items as structure location and test type have to be manually inserted into the test system.

To reduce the amount of labor involved in both the design and test of the structures used, JPL has developed a design and test program consisting of a Test Chip Assembler (TCA) and a Test Program Generator (TPG), which creates the geometrical description of the structures and generates the necessary test information using a high-level language. This system reduces the design time for a test chip by a factor of 30.

The Test Chip Assembler (TCA) consists of a set of Pascal modules containing the geometrical design rules and the procedures required to generate the geometrical description of the test structures. At the present time, this description is generated in the Caltech Intermediate Form (CIF). The individual structures are generated using parameterized procedures. The output of the TCA consists of:

(1) A CIF file, with the low-level geometrical structure description,

(2) A LST file, containing a human-readable description of the structures generated, and

(3) A TST file, containing the test information to be used by the Test Program Generator (TPG).

The Test Program Generator consists of a COMPILER program residing in a VAX 11-780 which executes the .TST file generated by the TCA. The digested code is then downloaded to an LSI 11/23 computer which controls the test system. A TRANSLATOR residing in the LSI 11/23 then transforms the program produced by the compiler into a form suitable for machine interpretation. An INTERPRETER in the test system then executes the translated code.

Both the TCA and the TPG are now operational.

[1]This section was prepared by C. A. Pina.

2.4.5 Statistical Analysis Package Capability[1]

To analyze the data obtained from wafer probing, a statistical package called STMJPL has been developed. It is based on a package called STAT2 [1] which is used at the National Bureau of Standards (NBS). As initially developed, STAT2 required very regular arrays in a wafer, with stringent requirements for the placement of structures for data analysis and wafer map generation. This made the program unsuitable for the irregularly placed test structures found in the foundry wafers used in the present program.

Some of the capabilities of the JPL software (STMJPL) are described below, especially those that make it different from STAT2. The package has been used successfully to analyze the data obtained from the DARPA-sponsored 1.25-μm CMOS bulk foundry fabrication runs.

2.4.5.1 Simultaneous Arrays. STMJPL works with up to twenty correlated arrays at a time. Therefore, if one thinks of the arrays as occupying adjacent columns in a matrix, then each row of that matrix represents a single data "point." For example, excluding an element in any given array causes the exclusion of all corresponding elements in the other arrays. Correlation statistics are available.

2.4.5.2 Position of Points on Wafer for Mapping Purposes. STMJPL allows data files to be read that contain as few as one data point, and will operate properly even if that single point becomes excluded. To generate wafer maps, as few as four points are needed, with the only requirement being that the points not all be colinear. STMJPL uses a smooth surface fitting algorithm from the computer science literature [2].

2.4.5.3 Recognition of Standard Structures and Relationship to Data Base. STMJPL will work properly whether or not the input file can be identified as "cross-bridge," "inverter," etc. However, code is in place that allows certain strictly formatted input files to be recognized by structure type. These raw data formats have long been in place at JPL. If a structure is recognized, then it is eligible for entry into the standard lot summary data base. In addition, certain structures, for example, the random-fault capacitor array, are such that computation of standard deviation is meaningless; however, STMJPL will automatically apply to the raw data a different analysis appropriate to that structure.

2.4.5.4 Value Wafer Maps. STMJPL permits a simple two-dimensional display, which is the outline of the wafer with the data points in position; each data point is represented by its value in floating-point format.

2.4.5.5 Character Wafer Maps. STMJPL provides some additional labeling and formatting of character wafer maps. The "streets" between chips are marked so that it is possible to know the relationship of the chip grid to the map. Scaling commands permit distortion to be minimized, and furthermore to be device-independent. It is possible to suppress all odd-valued digits, or alternatively all even-valued digits, to give "avenue" maps. The value scale includes the twelve characters: "-", "0", "1", ..., "9", "+".

[1]This section was prepared by C. A. Pina and G. A. Jennings.

2.4.5.6 Invertible Maps. STMJPL makes the initial assumption that "row 1" is at the bottom of the wafer; however, there is a simple command to tell STMJPL the opposite.

2.4.5.7 Contour Maps and Shaded Gray-Scale Maps. STMJPL does not yet have any capability to make contour maps or shaded gray-scale maps. It was found that gray-scale plots offered no significant visual improvement over the wafer maps produced by STMJPL. It is also not yet clear that there is any real value to contour map output. An independent three-dimensional plotting program was also developed (which displays "strata" of the independent variables in different colors), yet it offered little improvement to the visual information conveyed by STMJPL standard wafer maps.

2.4.5.8 Classification of Points. STMJPL works with three categories of points: included points (from which statistics are taken), excluded points, and non-points. Non-points occur when values are beyond reason, thereby representing a complete failure of the instrumentation to take a valid reading. Excluded points, on the other hand, are generally "real" readings that become excluded on the basis of knowledge of what is being measured, or statistical analysis.

2.4.5.9 Exclusion Trail Audit, and Printout. STMJPL keeps a detailed record of every command that might cause a change in status of a point. This record can be printed out on command. In addition, this can be stored as part of the data base.

2.4.5.10 Input Stream. STMJPL provides for command files, called "indirect files." In addition, STMJPL provides for true macro capabilities with textual parameter substitution. Indirect files and/or macros may invoke one another up to a nesting level of twenty. The STMJPL command processor provides for four types of command parameters, which are checked for type, number, and position: floating, integer, logical, and string parameters.

2.4.5.11 Array Processing. STMJPL provides facilities to modify the values of included and/or excluded points. These include multiplying by a power of ten, taking the logarithms of all points, and negating all points, on an array-by-array basis.

2.4.5.12 Summary Sheets. STMJPL contains the facility to produce summary sheets in predetermined formats for certain "recognized" structures. For example, there is a split-cross-bridge summary sheet and a transistor summary sheet. Although this code is structure-dependent, this is the most important data output format produced by JPL at this time.

2.4.5.13 Summary Sheet Reformatting. STMJPL provides for the post-processing of "summary sheets" that it produces. This is strictly format cleanup, such as the alignment of decimal points in columns, the elimination of columns of data, the exchange in position of columns of data, the "blanking" of redundant running label information, pagination cleanup, and sorting of rows. This utility will operate on a wide variety of tabular-format text files and could be invoked on many different types of summary sheets whether they are produced by STMJPL or not.

2.4.5.14 Summary Sheet Pre-Processing. STMJPL is capable of storing, in its lot summary data base, one entry of statistical information for each structure replicated over the wafer. The lot summary file is stored in this way. However, when the summary sheet is being prepared, it will often be found that two or more structures are identical in design. In this case the summary sheet can be "packed" by combining the data for the identical structures, recomputing the statistics, and writing a single entry to the summary sheet representing the combination of what would otherwise have been individual table entries. This also means there is a sorting capability, so that the summary sheet can be presented in an appropriate ordering of design parameters.

2.4.5.15 Log File. STMJPL contains a "log file" capability that can be enabled and disabled by command, and also provides for two levels of text inclusion to selectively capture output. STMJPL prints a count of generated errors and warnings when the command stream terminates.

2.4.5.16 Example STMJPL Character Wafer Maps. Some examples of STMJPL character wafer maps are shown in Figures 2.4.5-1 and 2.4.5-2. They are designed to fit onto single sheets for report purposes. The maps shown depict the distribution of values of a given parameter in the form of integers ranging from 1 through 9, where each integer represents an interval. The corresponding range of values of the intervals is given below each of the graphs.

2.4.5.17 References.

1. R. L. Mattis and R. Zucker, "Release Notes for STAT2 Version 1.31: An Addendum to NBS Special Publication 400-75," National Bureau of Standards Internal Report 83-2779 (November 1983).

2. H. Akima, "Bivariate Interpolation and Smooth Surface Fitting for Irregularly Distributed Data Points," Algorithm 526, ACM Transactions on Mathematical Software, Vol. 4, No. 2, pp. 160-164 (June 1978).

	MEAN	STDEV	% STDEV	MEDIAN	MINIMUM	MAXIMUM
RSHEET	37.84	1.83	4.83	37.99	34.48	39.92

INCLUDED = 12 EXCLUDED = 1 NON-POINTS = 1 N = 1.00

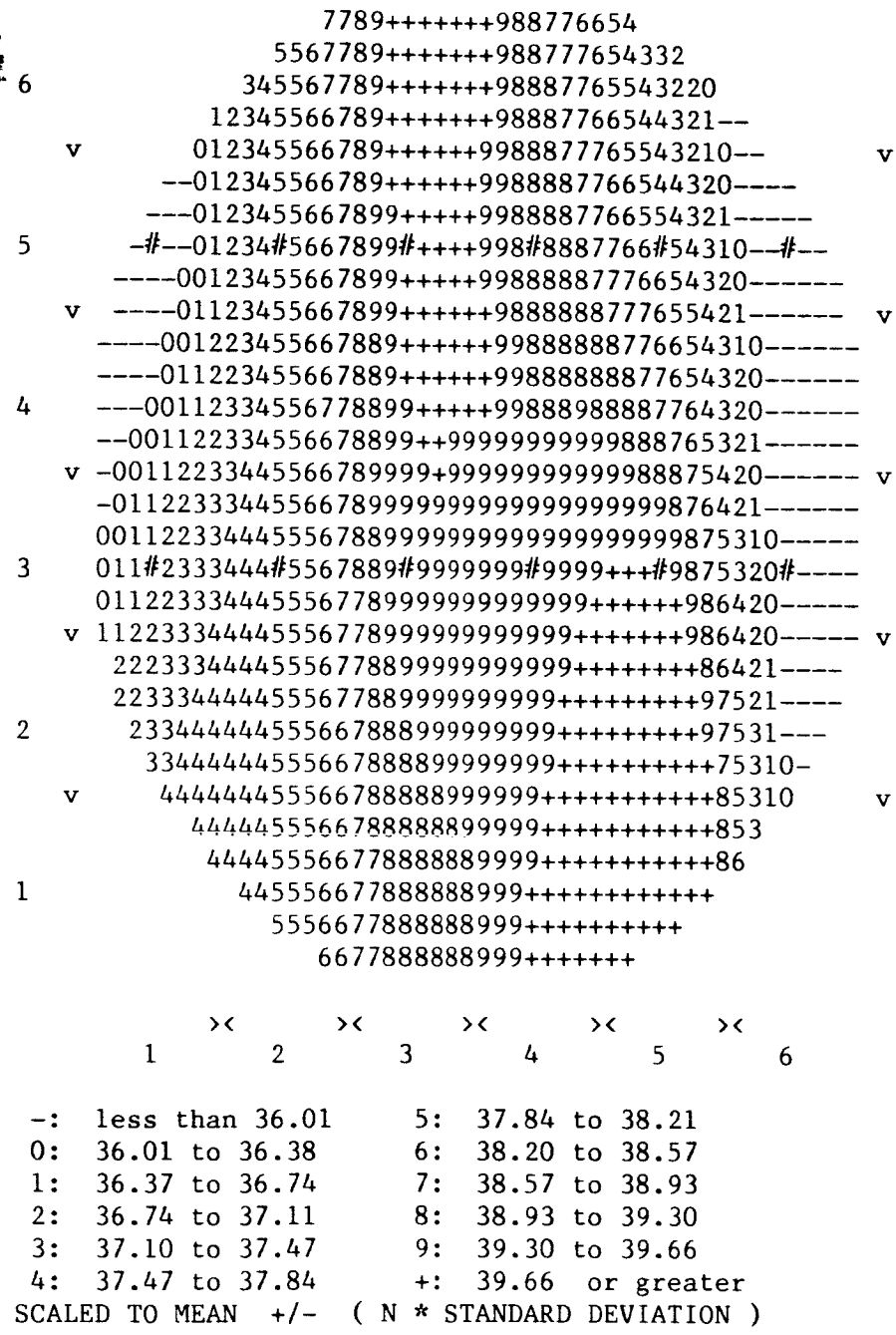


Figure 2.4.5-1. Sheet resistance (ohms/square) wafer map for n + poly layer measured from a split-cross-bridge resistor. The locations of 12 test structures on the 3-inch-diameter wafer are denoted by #. A smooth surface fitting algorithm [2] was used to generate the wafer map from the 12 measured sites.

```
          MEAN    STDEV    % STDEV   MEDIAN   MINIMUM   MAXIMUM
RSHEET    37.84   1.83       4.83    38.00     34.48     39.92

        INCLUDED = 12    EXCLUDED =  1    NON-POINTS =  1    N = 1.00

              ><       ><       ><       ><       ><

                      78899+++++9988877766
                    66778899+++++9988887766554
  6               5666778899+++++998888776655443
                345666778899+++++99888877766554321
    v         3345666778899+++++998888877665554321         v
            2234456667788999++++9988888777665543210-
          12234456667788999+++99988888777666543210-
  5       1#2234456#6778899#+++9998#8888777#6554321#--
        1122334456667778899+++99988888887776654321O---
    v   112233445666777889999+++99988888887776654321O--  v
      11222334456667778899999++99988888888777654321O---
      11223344456667778899999999988888888877654321O--
  4   12233344556667788899999999988888888877654321O--
      22333444556667788899999999988999988887654321O--
    v 23334445556667788899999999989999998887654321O-- v
      23344455556667778899999999999999999998876432210-
      33344455556666778899999999999999999999876533210-
  3   334#4555556#6677889#9999999#9999999#9886543#110-
      33444555566666778899999999999999999999875432110-
    v 34445555566666778889999999999999++++99875432210- v
       444555556666677888999999999999+++++9876432210
       445555556666677888899999999999++++++9976432210
  2     55555566666677888889999999999++++++97643221
        555566666667788888899999999++++++++9865322
    v    5566666666778888888899999++++++++986533      v
           666666667778888888999999+++++++++865
           66666667778888888899999++++++++++86
  1          666667778888888899999++++++++++
               6667778888888899999++++++++
                  777888888889999+++++

              ><       ><       ><       ><       ><
             1        2        3        4        5        6

   -:  less  than  3.4484E1        5:  3.7104E1 to 3.7651E1
   0:  3.4484E1 to 3.4992E1        6:  3.7651E1 to 3.8206E1
   1:  3.4992E1 to 3.5509E1        7:  3.8206E1 to 3.8770E1
   2:  3.5509E1 to 3.6033E1        8:  3.8770E1 to 3.9342E1
   3:  3.6033E1 to 3.6564E1        9:  3.9342E1 to 3.9923E1
   4:  3.6564E1 to 3.7104E1        +:  3.9923E1  or greater
```

Figure 2.4.5-2. Logarithmic wafer map for the data shown in Figure 2.4.5-1

2.5 FAULT MODELS

A major problem in the qualification of integrated circuit cells and in the development of adequate tests for the circuits is the lack of information on the nature and density of fault modes. Some of this information is being obtained from the test structures discussed in Section 2.3. In particular, the Pinhole Array Capacitor is providing values for the resistance of gate oxide shorts, and the Addressable Inverter Matrix is providing values for parameter distributions such as noise margins. In this section we examine another CMOS fault mode, that of the open-gated transistor, and attempt to assess the state of the transistors. That is, we wish to determine if an open-gated transistor is stuck on or stuck off. Preliminary results are described for a number of open-gated structures such as transistors, inverters, and NAND gates. In the latter part of this section, resistor faults are applied to various CMOS gates and the time responses for step inputs are noted using the circuit simulator SPICE. From these studies the critical value for the resistive short to upset the gate response was determined.

2.5.1 Open-Gate Transistors and Faulted Gates[1]

The purpose of this study is to determine the operating state of a transistor that has a floating gate that occurs during the manufacturing process. As shown in the inverter matrix test structure (Section 2.3.3), a floating gate can occur when contact to the gate is missing. To determine the transistor operating state, various faults were introduced into test structures consisting of arrays of faulted NAND gates, unfaulted inverters, and faulted inverters. Each of these structures used shift-register addressing to allow the evaluation of each device in the array. Results are presented in Tables 2.5.1-1, 2.5.1-2, and 2.5.1-3.

The faulted NAND gate array is composed of 104 good and 104 faulted gates arranged in a checkerboard pattern. A layout of the array is shown in Figure 2.5.1-1. Results are listed in Table 2.5.1-1. The faulted device had an intentionally open-gate connection to the p-channel pull-up transistor that would normally be connected to the n-channel transistor with a grounded source. The notation used in Table 2.5.1-1 is explained by the following example: VHIGH01 means that the V1 input is low, the V2 input high, and the output should be high for a "good" NAND gate. In the truth-table the faulted p-channel transistor is associated with the V2 input. To minimize the test sequence dependence, a 4-MΩ load resistor was placed between the output of the NAND gate and ground for the VHIGHA10 and VHIGHB10 tests. This load resistor caused the output of the good NAND gates to be degraded slightly from 5.02 V to 4.97 V.

The results in Table 2.5.1-1 show that the faulted NAND gates were not able to hold the output high for VHIGHA10 and VHIGHB10 tests. As seen in Table 2.5.1-1, the output was between 0.746 and 0.956 V. Taking an average of 0.85 V, the faulted p-channel transistor is passing a current of (0.85 V)/(4 MΩ) = 0.2 mA. Thus the pull-up transistor is partially on.

Results for an unfaulted inverter array are shown in Table 2.5.1-2, and these results should be compared with the results shown in Table 2.5.1-3. In Table 2.5.1-2, the results show that the inverters are pulling hard to the rails; that is, the output voltage is within 1 millivolt of either VDD or ground. Also, the inverter threshold voltage, VINV, is close to VDD/2, being about 2.3 V.

Results for the faulted inverter array are shown in Table 2.5.1-3. This array consists of 112 unfaulted inverters and 111 faulted inverters. The fault consisted of the omission of the poly link that connects the inverter to the poly input bus. The poly link that connects the pull-up and pull-down transistors remained intact. As seen in Table 2.5.1-3, the unfaulted inverters are not pulling hard to VDD because the faulted inverters are drawing current for the VHIGH test; that is, VDD = 4.94 V instead of 5.00 V as seen in Table 2.5.1-2. The faulted inverter measurements indicate that their outputs are stuck-at about 1.52 V for chip 2 and stuck-at 1.75 V for chip 3. This result can be explained if we assume that both the pull-up and pull-down transistors are partially on. A more detailed explanation is being pursued.

[1]This section was prepared by H. R. Sayah and M. G. Buehler.

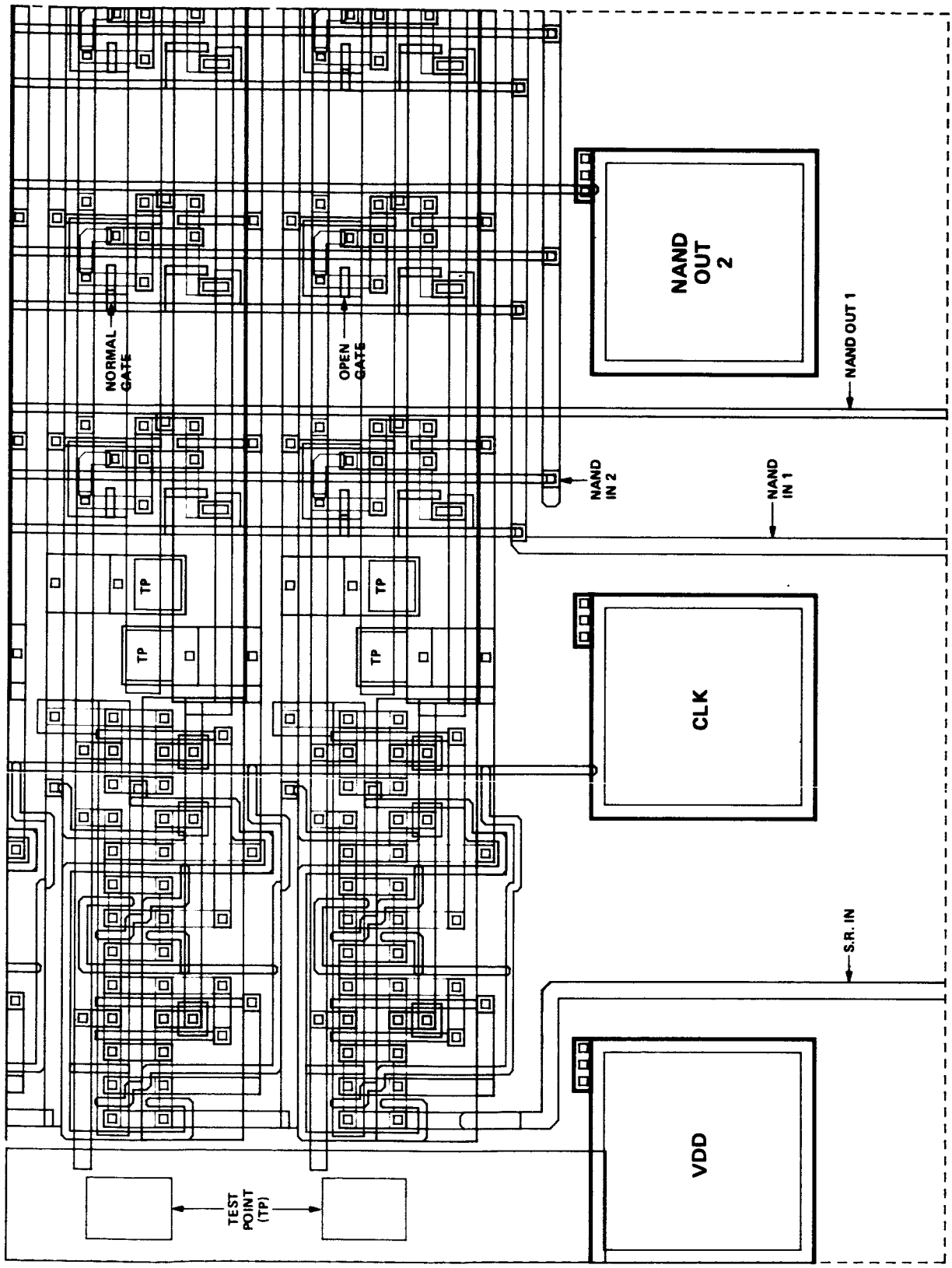


Figure 2.5.1-1. Layout of the addressable NAND gate matrix, a normal gate, and a gate with an open pull-up transistor

Table 2.5.1-1. 3-μm CMOS Bulk (p-Well) Faulted NAND Gate Array (4062/M46M/Chip 3)

TRUTH TABLE FOR GOOD AND FAULTY NAND GATES[a]

	V1	V2	VOUT(GOOD)	VOUT(FAULTED)
VHIGH00	L	L	H	H
VHIGHA10	H	L	H	L
VLOW11	H	H	L	L
VHIGHB10	H	L	H	L
VHIGH01	L	H	H	H

[a]Open-gate p-transistor connected to V2.

TEST RESULTS FROM UNFAULTED NAND GATES

	MEAN	SIGMA x 10^{-4}	PSIGMA	MEDIAN	MAXIMUM	MINIMUM	POINTS
VHIGH00	5.02	5.49	0.0110	5.02	5.02	5.02	104
VHIGHA10	4.97	43.8	0.0880	4.97	4.98	4.96	104.
VLOW11	0.00416	0.629	1.51	0.00419	0.00424	0.00399	104.
VHIGHB10	4.97	43.8	0.0881	4.97	4.98	4.96	104.
VHIGH01	5.02	5.55	0.0111	5.02	5.02	5.01	104
VINV	2.90	189.0	0.650	2.90	2.95	2.85	104.
GAIN	-24.69	5949.0	-2.41	-24.75	-22.53	-25.92	104.

TEST RESULTS FROM FAULTED NAND GATES[b]

	MEAN	SIGMA x 10^{-4}	PSIGMA	MEDIAN	MAXIMUM	MINIMUM	POINTS
VHIGH00	5.02	5.22	0.0104	5.02	5.02	5.02	104.
VHIGHA10	0.956	848.0	8.88	0.966	1.15	0.571	104.
VLOW11	0.0144	42.8	29.68	0.0136	0.0267	0.00401	104.
VHIGHB10	0.746	1329.0	17.82	0.714	1.11	0.498	104.
VHIGH01	5.02	5.34	0.0107	5.02	5.02	5.01	104.
VINV	2.63	186.0	0.706	2.63	2.68	2.58	104.
GAIN	-24.90	6492.0	-2.61	-24.84	-23.42	-26.30	104.

[b]Voltages VHIGHA10 and VHIGHB10 are measured for an output load of value 4 megohms.

Table 2.5.1-2. 3-μm CMOS Bulk (p-Well) Unfaulted Inverter Array (4062/M46M)

	MEAN	SIGMA x 10^{-4}	PSIGMA	MEDIAN	MAXIMUM	MINIMUM	POINTS	CHIP
VHIGH	5.00	0.281	a	5.00	5.00	5.00	222.	2
VLOW	a	0.105	23.9	a	a	a	222.	2
VINV	2.30	219.0	0.949	2.30	2.37	2.17	222.	2
GAIN	-23.25	8120.0	-3.49	-23.28	-21.04	-24.98	222.	2
VHIGH	5.00	0.208	a	5.00	5.00	5.00	222.	3
VLOW	a	a	25.2	a	a	a	222.	3
VINV	2.29	189.0	0.822	2.29	2.35	2.25	222.	3
GAIN	-23.20	7085.0	-3.05	-23.28	-20.66	-25.02	222.	3

[a]is equal to or less than 0.001.

Table 2.5.1-3. 3-μm CMOS Bulk (p-Well) Partially Faulted Inverter Array (4062/M46M)

TEST RESULTS FROM UNFAULTED INVERTERS						
	MEAN	SIGMA x 10^{-4}	PSIGMA	MEDIAN	POINTS	CHIP
VHIGH	4.94	60.5	0.123	4.94	112.	2
VLOW	0.00576	4.68	8.12	0.00583	112.	2
VINV	2.30	258.0	1.12	2.30	112.	2
GAIN	-22.56	7116.0	-3.15	-22.63	112.	2
VHIGH	4.94	60.3	0.122	4.94	112.	3
VLOW	0.00726	4.85	6.68	0.00733	112.	3
VINV	2.30	222.0	0.966	2.29	112.	3
GAIN	-22.43	10865.0	-4.84	-22.37	112.	3

TEST RESULTS FROM FAULTED INVERTERS								
	MEAN	SIGMA x 10^{-4}	PSIGMA	MEDIAN	MAXIMUM	MINIMUM	POINTS	CHIP
VHIGH	1.51	1690.0	11.2	1.49	1.89	1.17	111.	2
VLOW	1.53	1636.0	10.7	1.51	1.91	1.21	111.	2
VINV	1.50	1819.0	12.1	1.48	2.16	1.12	111.	2
GAIN	-0.623	2059.0	-33.0	-0.587	-0.168	-1.19	111.	2
VHIGH	1.75	1618.0	9.25	1.76	2.22	1.38	111.	3
VLOW	1.75	1609.0	9.17	1.76	2.23	1.39	111.	3
VINV	1.73	1577.0	9.12	1.73	2.19	1.35	111.	3
GAIN	-0.378	655.0	-17.3	-0.385	-0.177	-0.545	111.	3

2.5.2 MOS Integrated Circuit Fault Modeling[1]

2.5.2.1 Introduction. The faults considered here include only hard or permanent faults; soft faults or transient faults are not considered. Integrated circuit fault models have several uses: generation of test vectors for part screening, evaluation of field failure, and selection of fault tolerance methods. In this discussion, and consistent with generally accepted terminology, a failure is defined as any physical mechanism that results in permanent circuit damage. Furthermore, a failure causes a fault to occur. A fault is defined as the logical effect of a failure, and a fault can cause zero, one, two, or more system errors. An error is the manifestation of a fault on a single line logic value. For example, a transistor failure may cause the input of a gate to become stuck-at-zero. That faulty input will cause an error when an input value of one is applied to the faulty input and will yield no error if the input has value zero.

Fault models may be incorporated into simulation software for detailed and exhaustive analysis or may be used manually for quick look purposes. To be useful in a wide variety of applications, a fault model should have the following features:

- Low overhead: The model should not require awkward or excessive additions or modifications to the nominal circuit model.
- Compatible notation: Faults should be expressible in a notation compatible with that being used to model nominal behavior.
- Realistic fault set: The model must have the ability to accurately simulate the behavior of all integrated circuit failures.
- High level description: The model should be high level such that functional modeling of complex circuits is possible.

This report will look at three digital simulation techniques as well as physical simulation in modeling faults. These techniques embody a hierarchy of complexity bracketing the range of simulation levels. The digital simulation approaches are: transistor-level, connector-switch-attenuator level, and gate level. Physical simulation requires implementing actual integrated circuits into which failures have been inserted. This has the advantage of producing very precise fault modeling and measurement but has the disadvantage of having the highest cost and longest turnaround time of all the simulation methods. The lowest level digital model simulates fault behavior at the transistor level. It has the advantage of predicting circuit behavior with a high degree of accuracy. It has the disadvantages of requiring large computational resources for most circuits of interest, requiring failure specification at the circuit component or transistor parameter level, and requiring considerable human interpretation to evaluate the faulty behavior at

[1]This section was prepared by M. Sievers.

the logic level. At the other end of the digital simulation spectrum are the gate-level "work-around" models that permit simulation of faults at the gate level. These models have the advantage of producing relatively high level simulation and fault insertion but have the disadvantages of incompletely modeling integrated circuit failures and requiring complex networks to simulate faults. Connector-switch-attenuator digital simulation falls between the latter two models. This form of modeling provides accurate simulation of circuit behavior, is of high enough level to permit simulation of meaningful circuits, and provides easily interpreted results. The primary disadvantage of this modeling technique is that it is relatively new and consequently is not yet supported by simulation software.

Before closer examination of the modeling approaches described generically above, it is useful to examine the failure classes and associated faults.

2.5.2.2 Failure Characteristics. Integrated circuit reliability is affected by four broad mechanisms: manufacturing defects, random failures, wear-out, and externally induced failures. Manufacturing defects are flaws introduced by fabrication and handling processes. Random failures occur for various reasons during the useful life of the integrated circuit. Wear-out mechanisms, including metal migration and hot carrier injection, are those mechanisms that cause failures after many hours of operation. Excessive humidity, mechanical shock, radiation, electrostatic discharge, and excessive temperatures are examples of external failure mechanisms.

Manufacturing defects occur with different distributions which depend on the difficulty in performing a given process step. Results reported in [1] are typical for CMOS:

A. Photolithography errors

1. Oxide pinholes: $2.5/cm^2$
2. Missing features: $2.5/cm^2$
3. Source-drain shorts: $2.5/cm^2$

B. Metal errors

1. Metal shorts: $2.0/cm^2$
2. Metal opens: $2.0/cm^2$

C. Diffusion errors: $2.0/cm^2$

Random and wear-out related failures also have varying probability of occurrence. Results reported in [2] for CMOS are:

1. Stuck-on transistors: 35%
2. Gate oxide short: 25%
3. Metal failure: 20%
4. Quality related (scratches, bonding, etc.): 15%
5. Moisture leakage: 5%

These mechanisms induce five fault types: stuck-at faults, opens, coupling faults, timing degradation, and signal attenuation. From these data, stuck-at faults account for approximately 40% of all faults. Allocation of the remaining 60% depends on circuit details, but generally coupling faults and opens occur with lower probabilities than stuck-at faults but with higher probability than timing degradations and signal attenuation.

2.5.2.2.1 Stuck-At Faults. Stuck-at faults, comprising stuck-at-one and stuck-at-zero, are those faults that result in a signal line remaining high or low, respectively, during the time that the fault is present. These faults are the easiest to model, lead to well-defined behavior, and have been the traditional fault set of choice for research in testability and fault-tolerance. Many fault simulators exist for the stuck-at fault model. These simulators permit evaluation of complex digital circuits.

The exact percentage of all integrated circuit failure mechanisms resulting in stuck-at faults depends on numerous layout, manufacturing, and environmental parameters. It is known that stuck-at faults are not always the dominant fault type [1, 3, 4, 5]. As a result, these faults are a necessary component of an integrated circuit fault model but are not sufficient for a full characterization of the fault set.

2.5.2.2.2 Opens. Open faults are often referred to as stuck-at-X faults where X may be zero, may be one, or may vary between the two. This fault is caused by component or conductor failures that result in floating signal lines. The most dangerous effect of these faults is the creation of a parasitic storage element that increases the number of states in the logical machine [6]. For example, this fault will cause a combinational circuit to exhibit sequential behavior. The danger of this fault is that it can mask or delay error detection [7].

2.5.2.2.3 Coupling Faults. Coupling faults occur whenever signal lines designed to be independent in the fault-free circuit develop an influence on each other. This influence may be the result of capacitive coupling between signal lines or may be a direct connection through pin-holes in the insulating layers. Coupling faults cause a variety of behaviors, including stuck-at faults; AND, OR of the affected signals; pattern sensitivities; and voltages in the band between the zero and one voltages (stuck-at-X).

2.5.2.2.4 Timing Degradation. Elements within a logic circuit may suffer timing degradation due to several effects, including total dose radiation and hot-carrier injection. Observable disruption of normal operation occurs when the degradation has overrun a timing margin built into a circuit. Consequently, the effects of timing degradations become more apparent as the speed of the circuit increases.

2.5.2.2.5 Signal Attenuation. Several failure mechanisms result in attenuated voltage levels. These failures may result in stuck-at-one, stuck-at-zero, or stuck-at-X types of conditions. Asymmetric behavior of bilateral switches, i.e., the inability of the switch to correctly pass one voltage level due to the failure of a pass transistor, is another result of signal attenuation.

2.5.2.3 Physical Simulation Models. The failures indicated in Section 2.5.2.2 may be physically modeled by implementing collections of test circuits. These models contain or omit circuitry in order to simulate failure effects. At the analog circuit level, these failure effects may be simulated by missing connections, high impedance connections, extra connections, low impedance extra connections, higher impedance extra connections, and weak capacitive coupling to unrelated signals.

A simple set of circuits comprising inverters, NOR gates, NAND gates, and static and dynamic flip-flops may be used as the baseline functional elements onto which the above simulated failures are mapped. These circuits are illustrated in Figures 2.5.2-1 through 2.5.2-5. Circuit level failure simulation results for these circuits are found in Appendixes 2.5.2.A through 2.5.2.E. These simulations are based on the MICRO-CAP simulator and used the following parameters for the NMOS and PMOS transistors:

KP(W/2L)	=	0.01 mA/V^2	Conductance factor
VT	=	1 V	Threshold voltage
CGD	=	1E-11 F	Gate-to-drain capacitance
CGS	=	1E-11 F	Gate-to-source capacitance
RDS	=	10000 Ω	Drain-source on resistance
CDS	=	1E-11 F	Drain or source-to-body capacitance
GM	=	0.08 mS^{-1}	Transconductance

The failure simulations corroborate the failure models presented in the previous section. Stuck-at faults are often characterized by time-varying voltage waveforms that fail to traverse the 0/1 decision boundary. Timing degradations and signal attenuations are observed to be both symmetric and asymmetric. A symmetric failure is one that equally affects both the 0 to 1 transition and the 1 to 0 transition. An asymmetric failure affects only one of these transitions.

2.5.2.4 Digital Modeling. There are three levels at which modeling of integrated circuit failures is practical: the analog circuit, connector-switch-attenuator, and gate. Although it is possible to implement device physics level models, this level is too detailed to be useful for complex circuits. Modeling at the system level in which components are themselves complex circuits is too high-level to model detailed failure effects.

2.5.2.4.1 Transistor-Level Failure Models. Transistor-level modeling, as represented by such analysis programs as SPICE, permits evaluation of detailed circuit or component failures. In these models the circuit to be simulated is represented by a collection of interconnected electronic components (transistors, diodes, resistors, current sources, capacitors, and so forth). Semiconductor models are often highly detailed, permitting adjustment of many internal parameters.

A circuit to be modeled is treated as an analog circuit in the sense that it is modeled by time-dependent mathematical expressions. These expressions are solved to establish transistor bias conditions, nodal voltages, and so forth. Results obtained from these analyses are often tabulated or plotted as functions of time. The goal is to provide outputs resembling those derived from probing a physical implementation of that circuit with an oscilloscope. As with oscilloscope probing, the resulting time-voltage functions must be interpreted with respect to logical value representations and behavior.

Failures may be inserted into the circuit model in two ways. A large number of failure mechanisms may be simulated by inserting or deleting circuit elements. For example, an additional resistor could be included between two nodes to simulate coupling faults. The second mechanism requires modifying the characteristic parameters of the semiconductor devices. An example of this could be a threshold voltage to simulate the effect of radiation.

The mathematical analysis noted above is repeated after modifying the circuit or its parameter values. The nodal time functions can then be compared with the time functions of the fault-free circuit to characterize the effect of the fault.

2.5.2.4.2 Connector-Switch-Attenuator Level Failure Models. The connector-switch-attenuator (CSA) level of failure modeling simulates circuit operation at a higher level than the analog circuit level [8]. In this model, the circuit is represented by a collection of switches, conduction paths, and attenuators. Switches have positive and negative control inputs corresponding to NMOS and PMOS transistors, respectively. Loading is represented by attenuators.

Signals are constrained to discrete values; however, the model permits any number of values. The smallest set of values needed to describe a circuit in this model is four. These are 0, 1, high impedance (Z), and indeterminate or unknown (U). An additional modifier classifies a signal as being either strong or weak corresponding to relative current sourcing or sinking capabilities.

As an example, suppose that a particular circuit has k signals applied to a connector c. The connector c assumes an output value determined by the connection operation applied over the k signals. For example, if all but one signal had the value Z, c would have the value 1 if the remaining input had the value 1. If an equal number of signals had value 0 and 1, the output would be U. On the other hand, if one signal had the value 0 and all the remaining signals had the value 1~ (weak 1), the value of c would be 0.

A switch may be activated by either a strong or weak signal. It may be used to switch a strong signal. Switches therefore may be used to simulate the inherent signal amplification of gates. Switches may also be defined with asymmetric behavior. For example, a switch may attenuate a 1 to a 1~ but pass a 0 without attenuation.

A timing well is defined to enable simulation of time delay effects. The state of the well is determined by its initial conditions, the strength of the applied signals, the size of the well, and time. A timing well acts like a lumped resistor-capacitor (RC) time constant by delaying the charging or discharging of a connector. The nature of RC delays implies that strong signals will tend to make the well change state faster than weak signals.

An example of this model is illustrated in Figure 2.5.2-6. Figure 2.5.2-6a shows a conventional NMOS inverter characterized by a depletion load "pull-up" transistor and an enhancement mode "pull-down" transistor. An equivalent CSA model of the inverter is shown in Figure 2.5.2-6b. In the CSA

representation, A represents an attenuator, S is a switch, and W is a timing well. The inverter converts a 1~ to a 0 and a 0 to a 1~ to model the asymmetric behavior of NMOS circuits.

Analyzation of the circuit of Figure 2.5.2-6b follows traditional mechanisms for gate level models. The connectors are first set to initial signal values as determined by timing well or flip-flop initial conditions. The connector values are then evaluated by examining the signal values connected to each connector. Time is incremented and timing well outputs re-evaluated. Connector outputs are again computed.

Figure 2.5.2-6c shows the time behavior of Figure 2.5.2-6b. The input and output nodes x and z, respectively, take on values of 0 and 1~. By definition, a weak signal can source or sink less current than a strong signal. Consequently, the time required to alter the output by the weak input 1~ is longer than that required by the strong 0.

Modeling failures necessitate the addition/deletion of circuit elements or modification of attenuator, signal strength, or timing well values. For example, in Figure 2.5.2-6d an attenuator has been inserted between the switch and the output node. This simulates the effects of a faulty "pull-down" transistor. After modifying the circuit topology or timing parameters, circuit analysis proceeds as above.

2.5.2.4.3 Gate-Level Fault Models. Gate level modeling is the highest level modeling level suitable for fault studies of integrated circuits [6, 7, 9]. The classical model comprises gates, latches, flip-flops, and so forth. Signal levels are constrained to logical 1 and 0.

A circuit is modeled by its equivalent gate level logic diagram. The logic value of a connector is determined by the output driving that connector. Time behavior is determined in much the same manner as the previous model in that the value of each connector is determined as a function of time from the initial conditions.

Classical fault simulation at this level restricts the fault set to stuck-at-one and stuck-at-zero. This limitation has been moderated through the use of work-around models [6, 7]. In these models, complex logic circuits are substituted for the faulty device. This permits simulation of various opens and shorts. Timing behavior of the circuit elements may also be altered to simulate timing degradations.

Figure 2.5.2-7 shows the fault models for a CMOS NOR gate whose inputs are A and B and whose output is F. This figure can simulate the effects of stuck-at faults, open faults, and coupling faults. Components in this model indicated by an asterisk are assumed to fail in the classical stuck-at manner. The remaining components are assumed to be perfect. Labels near the fault components indicate the failure mechanism that causes the particular fault.

The original NOR gate is shown in the middle of Figure 2.5.2-7. It is marked with an asterisk to indicate that its output can faulted by a stuck-at-1 or stuck-at-0. The NOR gate output feeds the D input to a latch. The latch output follows the D input while the CLK control is 1. When CLK goes

to 0, the latch holds the output at the value of D when CLK was last 1. This latch is used to simulate parasitic storage elements that form when the NOR gate output is neither pulled to Vdd or ground. When a failure of this type occurs, the previous output of the NOR gate will be held in the gate's fan-out capacitance.

There are five failures that result in an open in the NOR output: B PULL-DOWN MISSING, A PULL-DOWN MISSING, BROKEN OUTPUT WIRE, MISSING OUTPUT CONTACT, and P TRANSISTOR DEFECTIVE. There is a buffer associated with each of these failures. The output of these buffers will be 0 if no failure is to be simulated and 1 otherwise. For example, if B PULL-DOWN MISSING is set to 1 and if input B = 0 and input A = 1, the CLK input to the latch will be 0. This condition simulates the effect that the NOR gate will float for this fault only when the NOR gate output depends on the defective transistor.

Coupling faults in the form of shorts (bridging) are modeled with a multiplexer created from two AND gates and an OR gate. In normal operation the multiplexer selects the nominal signal for its output. When a coupling fault is simulated, the multiplexer outputs the AND of the nominal signal and a neighboring signal.

For example, one such multiplexer is located above the original NOR gate. This multiplexer selects either the NOR gate output when no fault occurs or the AND of the NOR with NEIGHBOR 3 when a bridging fault is being simulated. NEIGHBOR 3 is a signal wire near the NOR gate output. In CMOS logic, the value 0 normally dominates the value 1. As a result, if two independent outputs short together, the result will be the AND of the two outputs. Bridging of signals NEIGHBOR 1 and NEIGHBOR 2 with the two inputs is simulated by the multiplexers located at the bottom of the figure.

2.5.2.4.4 Model Comparison. Each of the generic models described above is useful for modeling integrated circuit failures. The models vary primarily in the range of applications for which they are to be used. That is, when specifying a model it is important to determine the modeling requirements. Very detailed analysis of small circuits is best done by a different model than that used for crude analysis of very large circuits. The features, advantages, and disadvantages of each of the modeling techniques are summarized in the subsections below.

2.5.2.4.4.1 Transistor-Level Modeling.

Features:

- Detailed analysis at the electronic component level.
- Results are presented in a manner analogous to an oscilloscope probe of an actual implementation.
- Many computerized simulation packages are available.

Advantages:

- Highest accuracy in modeling circuit behavior.
- Permits modification of transistor behavior to simulate intra-transistor failures.

Disadvantages:

- Requires solution of generalized time-dependent differential circuit equations.
- Practical application range limited to analyzing small circuits.

2.5.2.4.4.2 Connector-Switch-Attenuator Level Modeling.

Features:

- Simplified representation of integrated circuit components.
- Results are in the form of a logic analyzer output.

Advantages:

- Straightforward means to model logic level effects of integrated circuit failures.
- Useful in modeling larger circuits than the transistor level model.

Disadvantages:

- Does not model failures with the same detail as transistor-level simulation.
- Simulation software is not yet generally available.

2.5.2.4.4.3 Gate Level Modeling.

Features:

- Primitives are gates, flip-flops, and latches.

Advantages:

- Models large networks effectively.
- Many software simulation packages are available.

Disadvantages:

- Does not model all relevant integrated circuit failures.
- Complex work-around models are needed to simulate faults other than the stuck-ats.

2.5.2.5 Conclusion. By their nature, digital components are less susceptible to small variations in supply voltage, transistor gains, capacitances, resistances, timing, and so forth than analog circuits. More spectacular upsets such as those that cause shorts, opens, significant timing or voltage level degradation, or stuck signal wires do alter circuit behavior in a measurable way.

Appendixes 2.5.2.A through 2.5.2.E provide a survey of the typical subtle and major circuit upsets. As evidenced by the simulations, failed circuits exhibit behaviors ranging from no apparent malfunction to asymmetric timing degradation or signal attenuation. The onset of erroneous behavior corresponds to values of contact resistance observed in test structures.

Based on the observed fault classes, the connector-switch-attenuator model appears to be the best fault simulation compromise of the three modeling methods. It remains to be demonstrated that this fault set fairly represents faults in actual circuits. A simulator is currently being developed but has not yet been released [10].

2.5.2.6 References.

1. G. R. Case, "Analysis of Fault Mechanisms in CMOS Logic Gates," Proc. 13th Design Automation Conference, 265-270 (June 1976).

2. G. L. Schnable, et al., "Reliability of CMOS Integrated Circuits," Computer, 11 (10), 6-17 (October 1978).

3. M. F. Adam and D. M. Aaron, "MOS Reliability Prediction Model," Annals of Reliability and Maintainability Conference, Vol. 9, 576-586 (July 1970).

4. E. S. Anolick and G. R. Nelson, "Low-Field Time-Dependent Dieletric Integrity," IEEE Trans. Reliab., R-29 (3), 217-220 (August 1980).

5. S. J. Rosenberg, D. L. Crook, and B. L. Euzent, "H-MOS Reliability," IEEE Trans. Electron Devices, ED-26 (1), 48-51 (January 1979).

6. R. L. Wadsack, "Fault Modeling and Simulation of CMOS and MOS Integrated Circuits," Bell System Tech. J., 57 (5), 1449-1473 (May-June 1978).

7. M. W. Sievers and A. Avizienis, "Analysis of a Class of Totally Self-Checking Functions Implemented in a MOS LSI General Logic Structure," Proc. Eleventh Annual International Symposium on Fault-Tolerant Computing, 256-261 (June 1981).

8. J. P. Hayes, "Fault Modeling for Digital MOS Integrated Circuits," IEEE Trans. Computer-Aided Design, CAD-3 (3), 200-207 (July 1984).

9. J. Galiay et al., "Physical Versus Logical Fault Models in MOS LSI Circuits, Impact on Their Testability," Proc. Ninth Annual International Symposium on Fault-Tolerant Computing, 195-202 (June 1979).

10. M. Kawai and J. P. Hayes, "An Experimental MOS Fault Simulation Program CSASIM," Proc. 21st Design Automation Conference (June 1984).

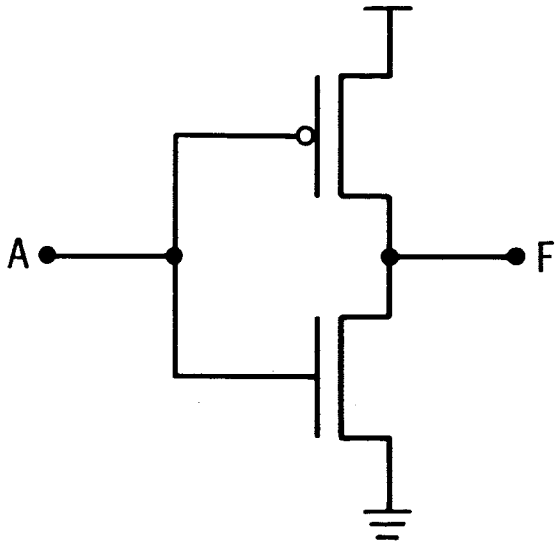


Figure 2.5.2-1. Inverter

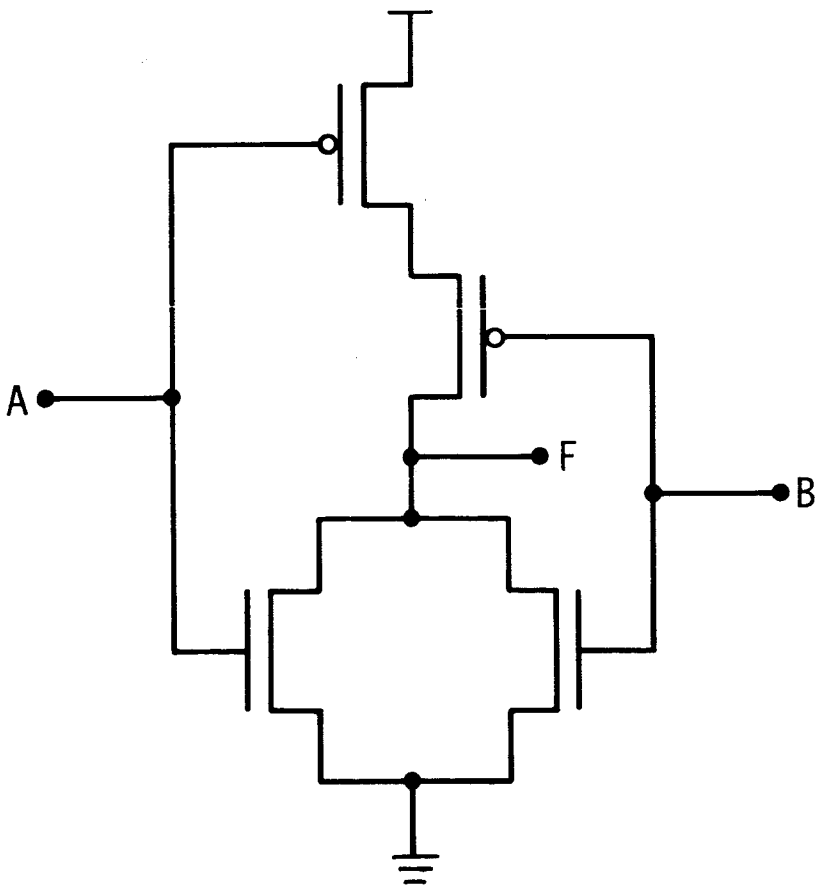


Figure 2.5.2-2. NOR gate

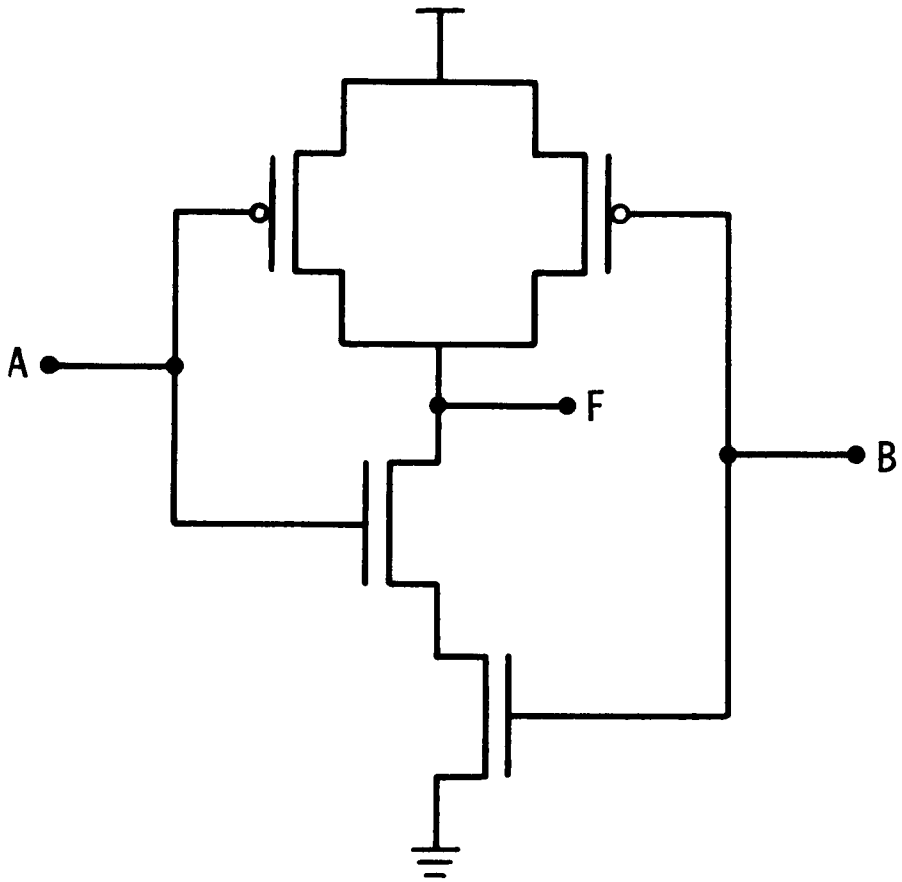


Figure 2.5.2-3. NAND gate

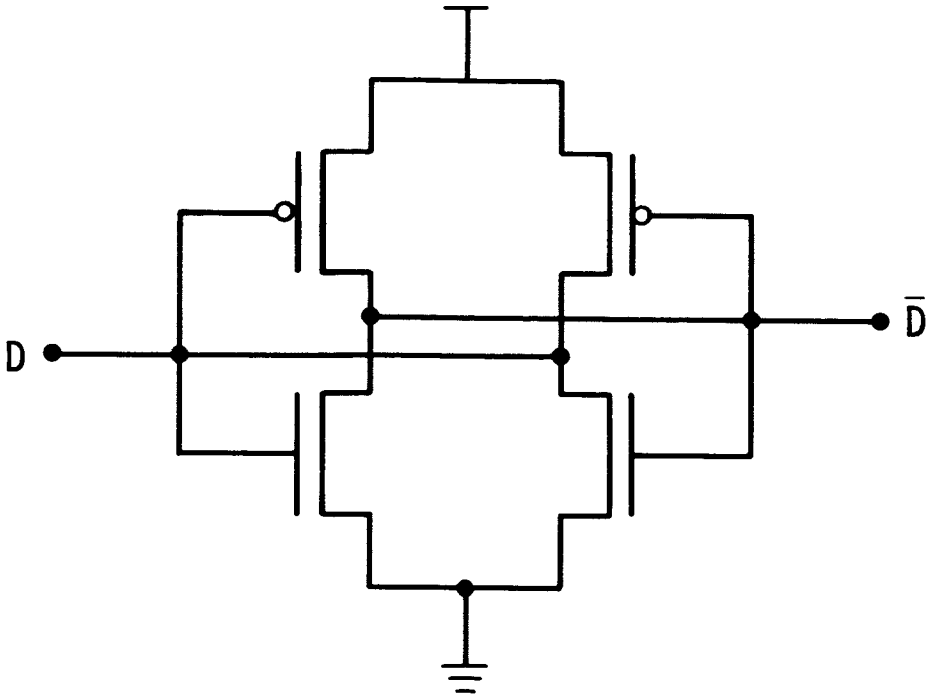


Figure 2.5.2-4. Static flip-flop

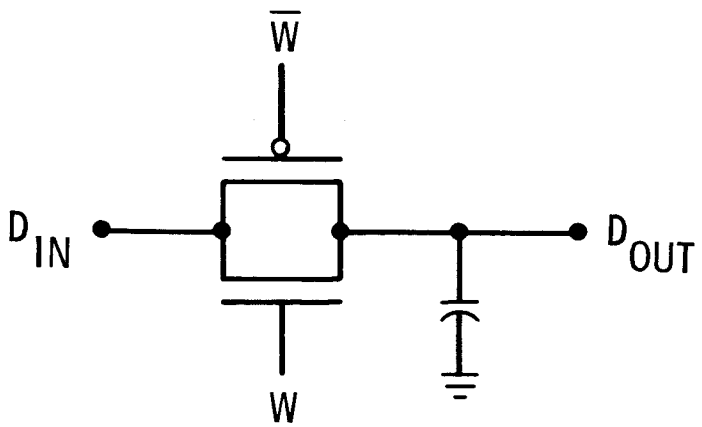


Figure 2.5.2-5. Dynamic flip-flop

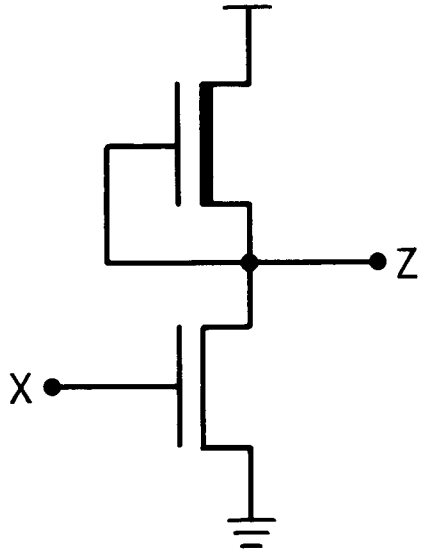


Figure 2.5.2-6a. NMOS inverter

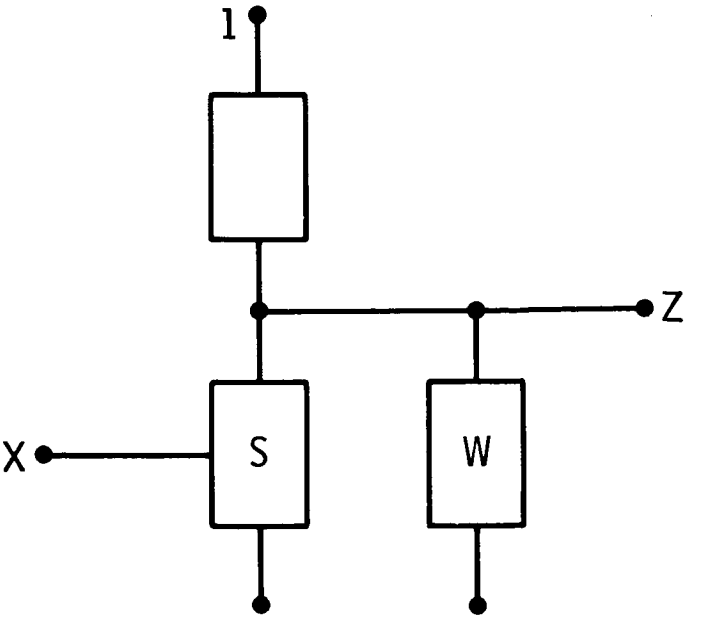


Figure 2.5.2-6b. CSA inverter model

NODE	X	Z
TIME		
0	0	1~
1	1~	1~
2	1~	1~
3	1~	1~
4	1~	0~
5	1~	0~
6	1~	0
7	1~	0
8	0	0
9	0	0~
10	0	1~
11	0	1~

Figure 2.5.2-6c. Inverter timing

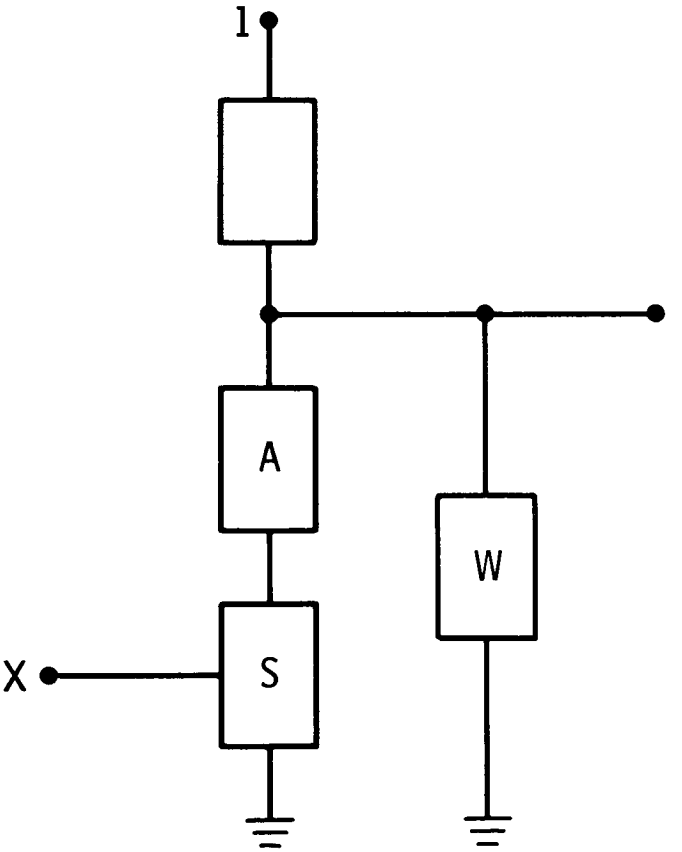


Figure 2.5.2-6d. Failed inverter

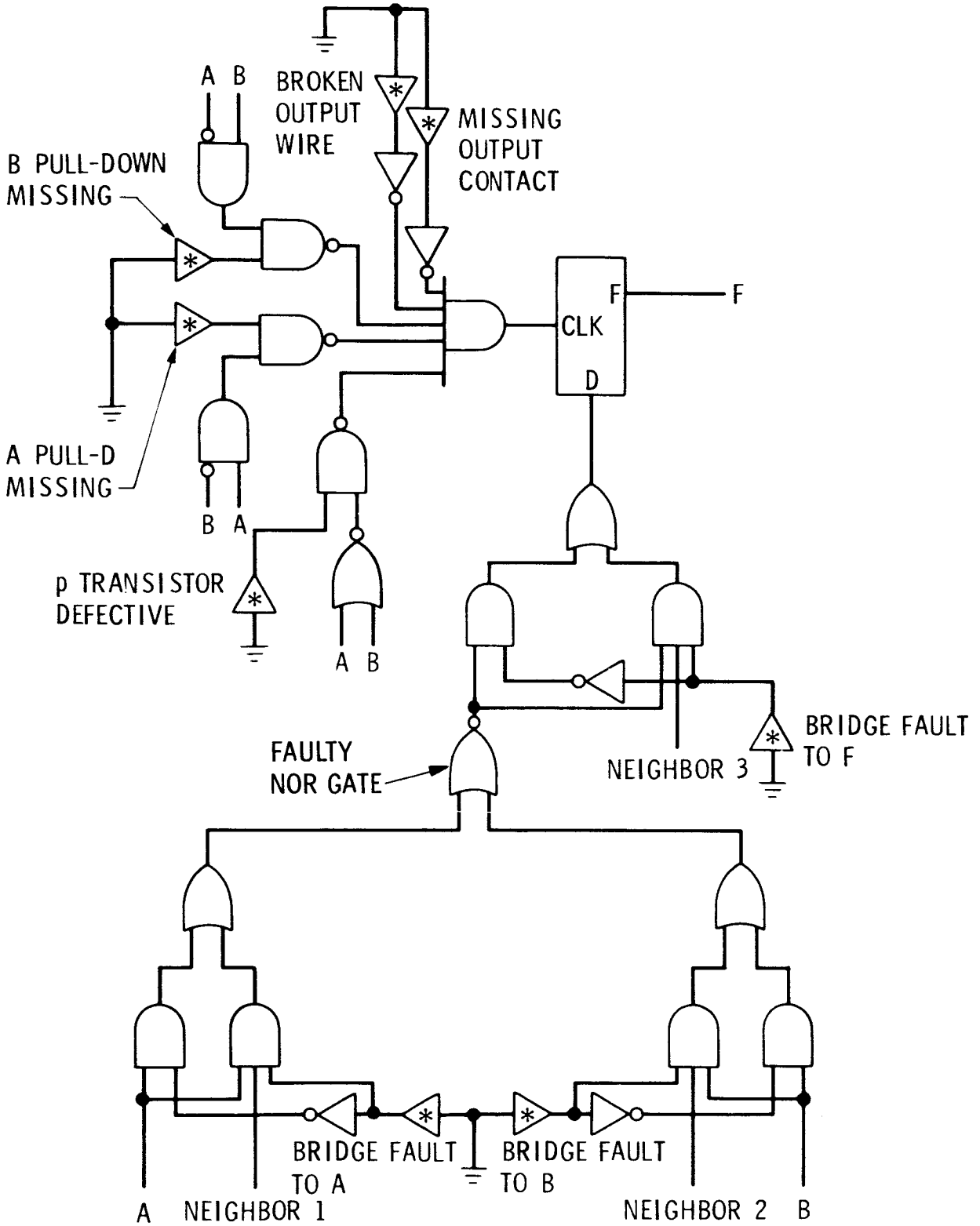


Figure 2.5.2-7. CMOS gate level work-around model for NOR gate

APPENDIX 2.5.2.A

CMOS INVERTER TRANSISTOR-LEVEL FAULT MODEL ANALYSIS

A CMOS inverter was analyzed using a transistor-level circuit analysis program running on the IBM PC computer. The 15 simulations run on the inverter illustrate the following behaviors: stuck-at-0, stuck-at-1, attenuated pull up, attenuated pull down, attenuated and asymmetric pull up and pull down, and timing degradations.

The fault model shown in Figure 2.5.2.A-1 utilizes two pulse sources, a capacitor, eight resistors, and a PMOS and NMOS transistor. Signal source S0 drives the inverter input. Source S1 models a nearby signal source that is capacitively coupled to the inverter output. The plot shows the outputs of S0 and S1 where S0 rises to 5 volts and S1 peaks at 3.5 volts. S1 was attenuated to simulate a drop in the coupled voltage. S1 pulses occur once during S0 at Vdd and once at S0 = 0 volts. The resistors and capacitor used to model various intra-inverter opens and shorts are defined as follows. Nominal resistor and capacitance values are listed below.

R1:	S0 to output	1E6 Ω
R2:	S0 to inverter input	0.1 Ω
R3:	S0 to P gate	0.1 Ω
R4:	P drain to N gate	0.1 Ω
R5:	N source to ground	0.1 Ω
R6:	Vdd to P source	0.1 Ω
R7:	P source-drain short	1E6 Ω
R8:	N source-drain short	1E6 Ω
C1:	Capacitor	1E-11 F

Figure 2.5.2.A-2 shows the normal output operation of the inverter. The lower trace in Figure 2.5.2.A-2 shows the expected inverter operation.

Figure 2.5.2.A-3 shows the inverter behavior when the input is shorted to the output. The source S0 has sufficient drive to dominate the inverter operation.

In Figure 2.5.2.A-4 the input/output short is 10 ohms. The inverter output still follows its input; however, the output is attenuated. The inverter behaves as though its output were stuck-at-1.

In Figure 2.5.2.A-5 the input/output short is 100 ohms. Here the inverter behavior dominates the short although the "zero" level has been raised to 0.3 volts and the "one" level has been lowered to 4.5 volts.

Figure 2.5.2.A-6 models a "moderate" open from S0 to the inverter by setting R2 to 1000 ohms. Here the NMOS transistor does not conduct. As seen, the resulting inverter output is stuck-at-1.

Figure 2.5.2.A-7 simulates a "moderate" open to the PMOS transistor by setting R3 to 1000 ohms. In this figure the P transistor fails to completely turn off when the input is at Vdd. The resulting output is produced by a voltage divider created from the source-drain resistances of the N and P transistors. Note that as the RC time constant in the P-transistor gate charges up, the inverter output drops when S0 goes to Vdd. When the input goes low, the P transistor does not conduct and bring the output to Vdd, so the output remains low.

Figure 2.5.2.A-8 simulates a 1-megohm "open" to the PMOS transistor by setting R3 to 1 megohm. The RC time constant at the P-transistor gate is large enough to prevent the inverter output from changing as it did in Figure 2.5.2.A-7 when S0 is at Vdd. The inverter goes to Vdd when the input goes to zero.

In Figure 2.5.2.A-9 a 1000-ohm resistance is placed between the drain of the P-transistor and the output. The inverter output is pulled partway to ground and exhibits an RC time constant on pulling up. This is the only simulation in which the effect of the coupled source is evident.

Figure 2.5.2.A-10 illustrates inverter behavior when a 1-megohm resistor separates the P-transistor from the output. This simulation shows that the inverter output fails to pull up to Vdd. Its operation is dominated by the coupled source S1 when the inverter input is zero. The coupled source causes a jump in excess of 1 volt when the inverter input is low and has no apparent effect when the N-transistor is conducting.

Figure 2.5.2.A-11 represents the inverter behavior with a 1000-ohm resistance between the N-transistor source and ground. In this simulation, the inverter output is stuck-at-1.

Figure 2.5.2.A-12 shows inverter behavior when a 100-ohm resistance is placed between the P-transistor source and Vdd. In this simulation, the inverter pulls down normally but exhibits an RC time constant delay on pulling up.

Figure 2.5.2.A-13 shows inverter behavior when there is a high resistance to Vdd. As would be expected, the inverter behaves stuck-at-0.

In Figure 2.5.2.A-14 the P-transistor source-drain has been shorted by a 10-ohm resistor. A stuck-at-1 behavior results.

Figure 2.5.2.A-15 is a simulation of a higher impedence short across the P-transistor. In this figure, the inverter appears to behave normally.

Finally, Figure 2.5.2.A-16 shows the simulation of a low impedence short across the N-transistor. Here the inverter pulls down normally but attenuates the pull up voltage.

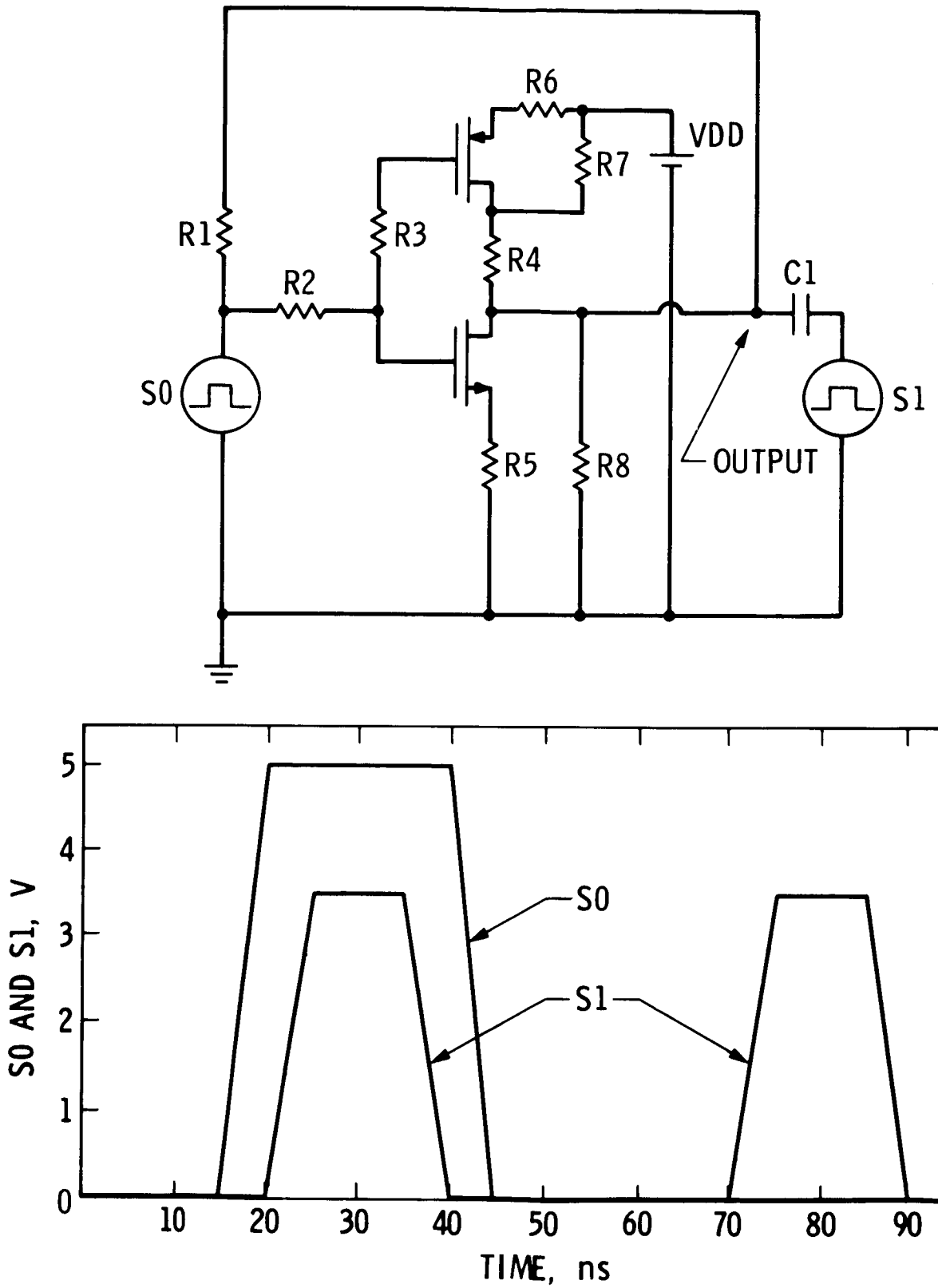


Figure 2.5.2.A-1. Inverter fault model

Figure 2.5.2.A-2. Nominal output behavior for an inverter fault model

Figure 2.5.2.A-3. R1 = 1 Ω for an inverter fault model

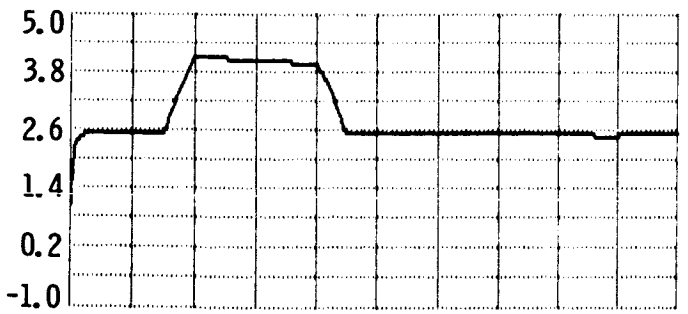


Figure 2.5.2.A-4. R1 = 10 Ω for an inverter fault model

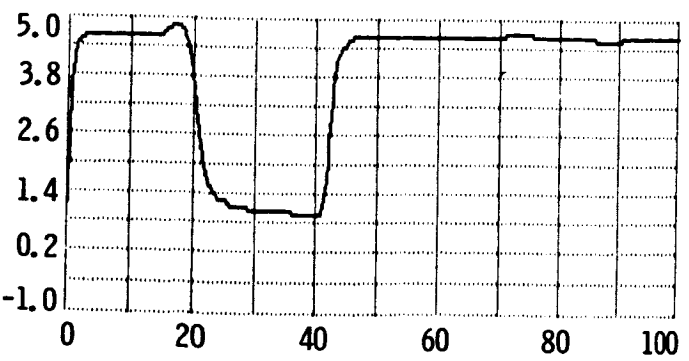


Figure 2.5.2.A-5. R1 = 100 Ω for an inverter fault model

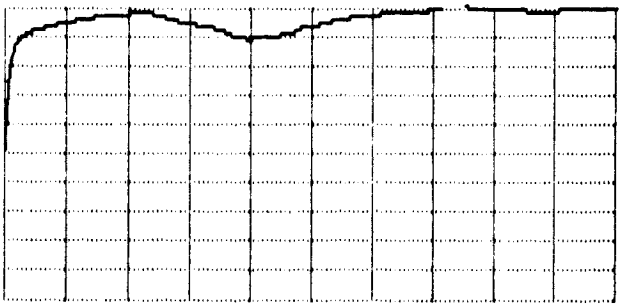


Figure 2.5.2.A-6. R2 = 1,000 Ω for an inverter fault model

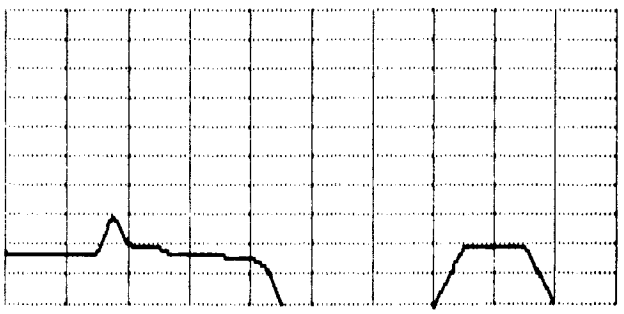


Figure 2.5.2.A-7. R3 = 1,000 Ω for an inverter fault model

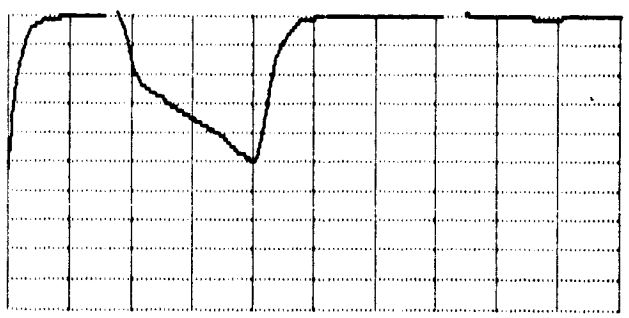


Figure 2.5.2.A-8. R3 = 1,000,000 Ω for an inverter fault model

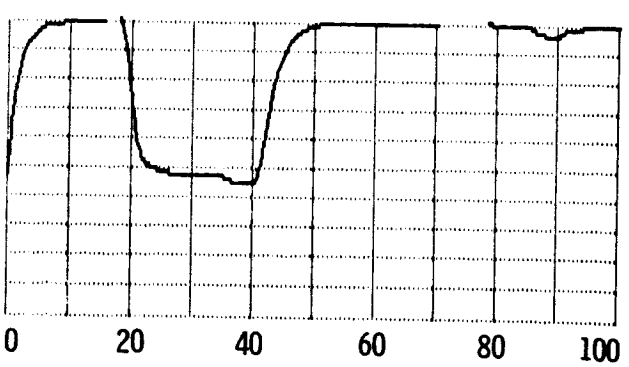


Figure 2.5.2.A-9. R4 = 1,000 Ω for an inverter fault model

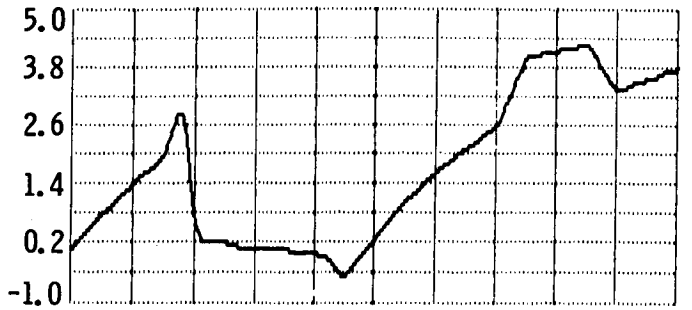


Figure 2.5.2.A-10. R4 = 1,000,000 Ω for an inverter fault model

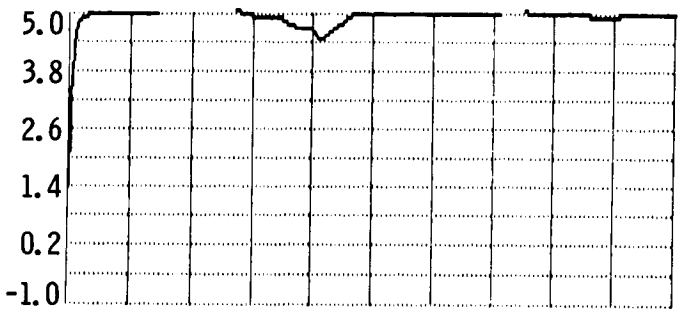


Figure 2.5.2.A-11. R5 = 1,000 Ω for an inverter fault model

Figure 2.5.2.A-12. R6 = 100 Ω for an inverter fault model

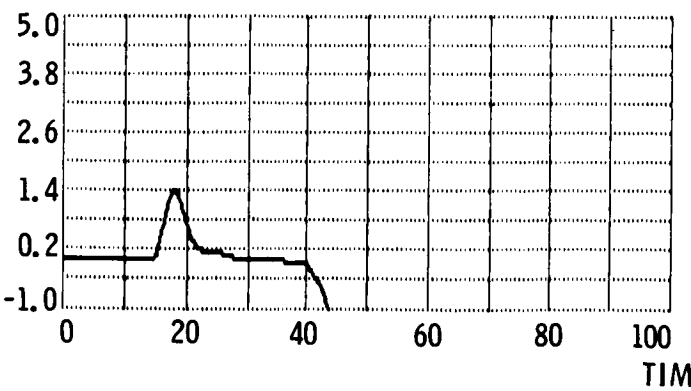


Figure 2.5.2.A-13. R6 = 1,000,000 Ω for an inverter fault model

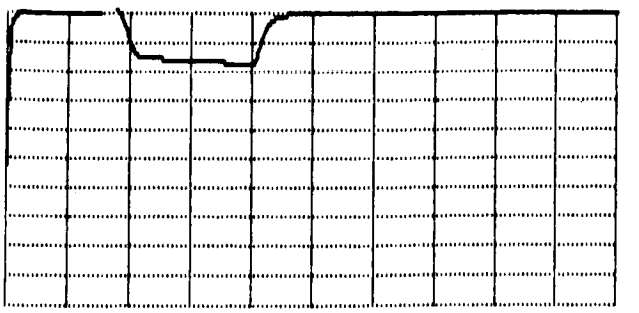


Figure 2.5.2.A-14. R7 = 10 Ω for an inverter fault model

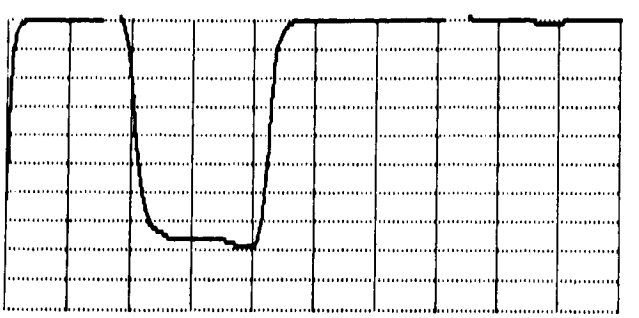


Figure 2.5.2.A-15. R7 = 1,000 Ω for an inverter fault model

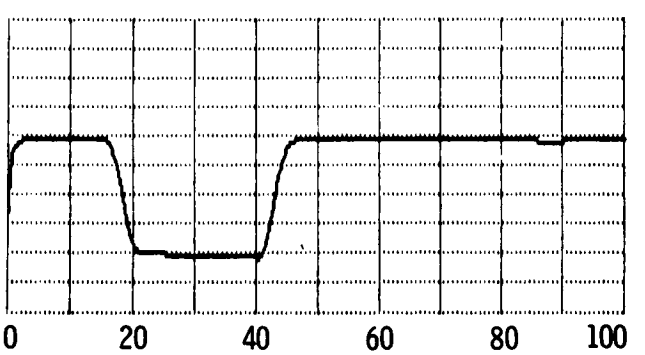


Figure 2.5.2.A-16. R8 = 10 Ω for an inverter fault model

APPENDIX 2.5.2.B

CMOS NAND TRANSISTOR-LEVEL FAULT MODEL ANALYSIS

A CMOS NAND gate was analyzed using a transistor-level analysis program. The simulations indicate that the NAND gate exhibits stuck-at as well as asymmetric timing and signal degradation effects.

The fault model shown in Figure 2.5.2.B-1 comprises three independent pulse sources, two P-type transistors, two N-type transistors, a coupling capacitor, and 15 resistors. Nominal resistor values are indicated below. The resistors and capacitor are defined as follows:

R1:	Open to ground	0.1 Ω
R3:	Open between N-transistors	0.1 Ω
R5:	Open from output to N-transistor	0.1 Ω
R6:	Open from output to P-transistor	0.1 Ω
R8:	Open from P-transistor to output	0.1 Ω
R10:	Open from Vdd to P-transistor	0.1 Ω
R11:	Open from input	0.1 Ω
R12:	Open from input to P-transistor	0.1 Ω
R13:	Short from input to Vdd	1E6 Ω
R14:	Open from input to N-transistor	0.1 Ω
R15:	Short from input to ground	1E6 Ω
R18:	Short from output to Vdd	1E6 Ω
R19:	Short from output to ground	1E6 Ω
R23:	Short from input to output	1E6 Ω
R24:	Open from Vdd	0.1 Ω
C1:	Capacitance	1E-11 F

The capacitor couples an attenuated signal source to the NAND output to simulate capacitive coupling to a nearby signal.

The traces in Figure 2.5.2.B-1 represent the inputs to the NAND gate. The 3.5-volt pulses in the remainder of the figures in this appendix represent the coupled signal. The nominal NAND gate output appears in Figure 2.5.2.B-2.

In Figure 2.5.2.B-3, resistor R1 was set to 100 ohms, simulating a moderate impedance open to ground. The resulting output waveform is the expected stuck-at-1.

Figure 2.5.2.B-4 simulates a moderate open between the N-transistors. The NAND gate output suffers a timing degradation and signal attenuation on pull down.

A moderate open is simulated in Figure 2.5.2.B-5 between the output and the N-transistors. Here the NAND gate exhibits a stuck-at-1 behavior.

Figure 2.5.2.B-6 simulates a moderate open to the pull up P-transistors. A timing degradation occurs in a 0 to 1 transition of the NAND gate output.

In Figure 2.5.2.B-7 the open from the output to the P-transistors was increased. Here the timing degradation to achieve a logic 1 value is significant and the NAND gate output becomes more sensitive to the coupled voltages when the ouput should be high.

Figure 2.5.2.B-8 shows a moderate short from the output to Vdd. There is a broadening of the negative pulse, indicating a timing degradation.

Figure 2.5.2.B-9 represents an open between the drains of the P-transistors. A timing degradation occurs in pulling the NAND gate output to Vdd due to the open to the leftmost P-transistor.

A moderate open from a P-transistor to Vdd is simulated in Figure 2.5.2.B-10. Again a timing degradation occurs.

In Figure 2.5.2.B-11, a moderate open is inserted between one of the signal sources and the NAND gate. Here, the NAND gate appears stuck-at-1.

Figure 2.5.2.B-12 simulates a moderate open to the gate of the leftmost P-transistor. The resulting output is attenuated on the low going pulse due to both the faulted P-transistor and both N-transistors conducting. This fault behaves as a signal degradation.

In Figure 2.5.2.B-13, a low impedence short connects a signal input to Vdd. The resulting waveform shows little change in the NAND gate output due to the high drive capability of the signal source.

Figure 2.5.2.B-14 simulates a moderate impedence open from the source to an N-transistor. The resulting NAND gate output suffers a timing degradation.

In Figure 2.5.2.B-15, a low impedence short is placed from ground to the gate of an N-transistor. Here the NAND gate exhibits both a timing degradation and a signal attenuation to ground.

Figure 2.5.2.B-16 shows a low impedence short from the output to Vdd. The resulting output is stuck-at-1.

In Figure 2.5.2.B-17, a low impedence short was placed from the output to ground. The NAND gate suffers a timing degradation on pulling up as well as a signal attenuation.

A low impedence short from input to output is simulated in Figure 2.5.2.B-18. Here the NAND gate appears stuck-at-1.

Figure 2.5.2.B-19 simulates a moderate open from Vdd to the P-transistors. The output waveform shows a significant timing degradation on pull up as well as an interaction with the coupled signal.

A smaller open from Vdd to the P-transistors is simulated in Figure 2.5.2.B-20. Here the NAND gate primarily exhibits a timing degradation on pulling up.

Finally, Figure 2.5.2.B-21 simulates an open from the output to the N-transistors. The coupled signal has a pronounced effect when the NAND gate output is floating. The NAND gate appears stuck-at-1.

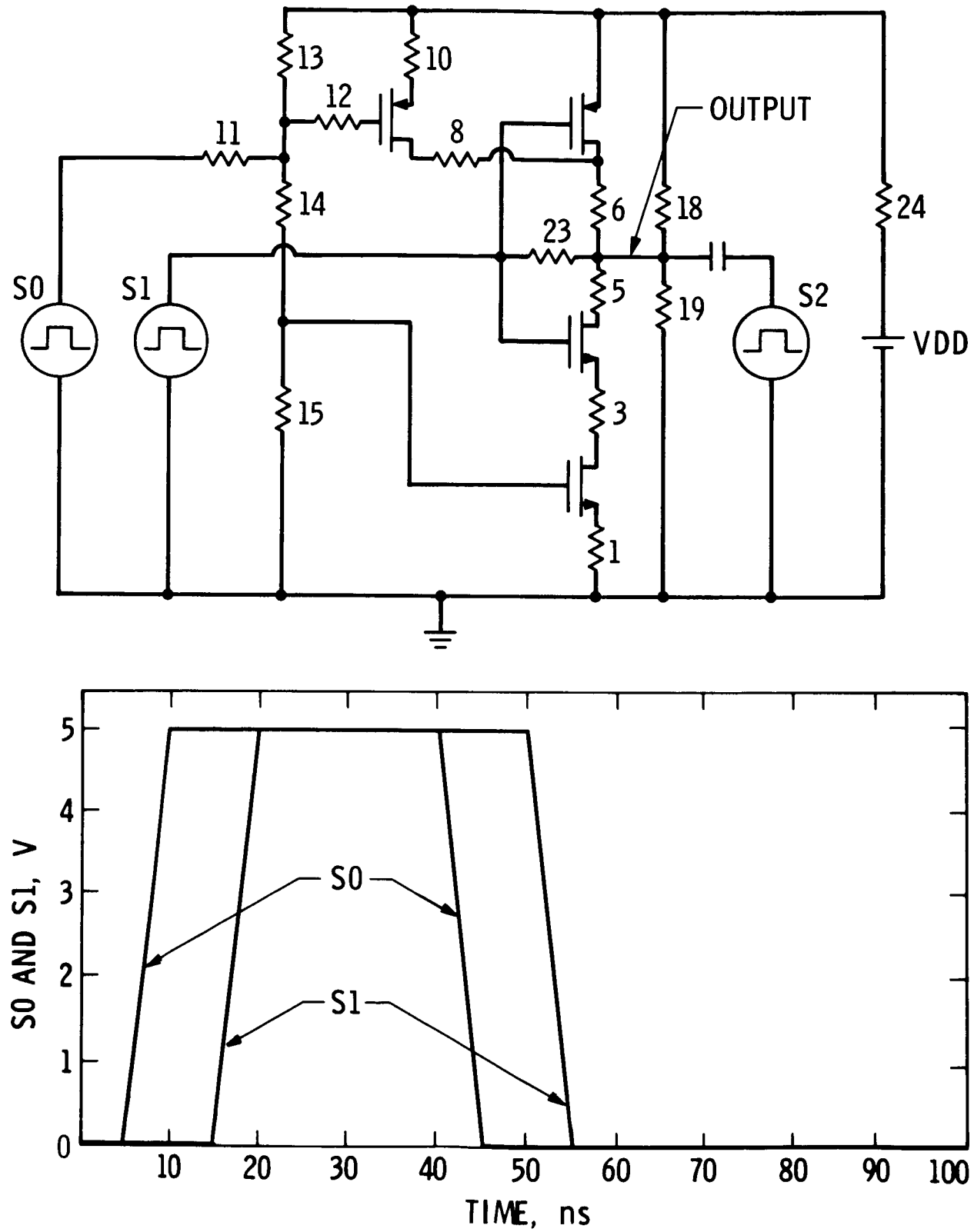


Figure 2.5.2.B-1. NAND fault model

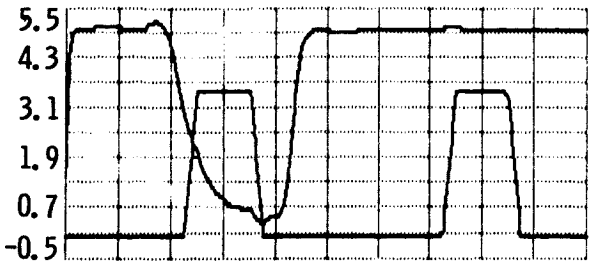


Figure 2.5.2.B-2. Nominal behavior for a NAND fault model

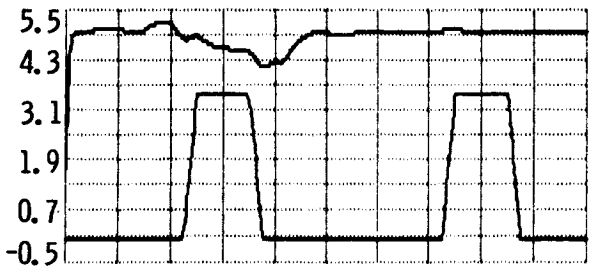


Figure 2.5.2.B-3. R1 = 100 Ω for a NAND fault model

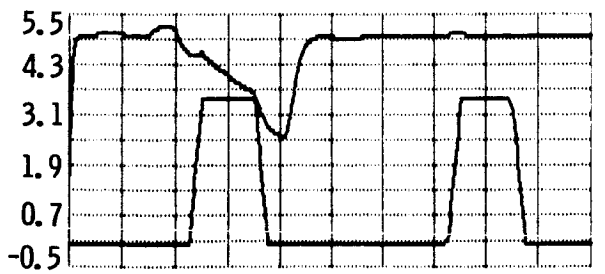


Figure 2.5.2.B-4. R3 = 50 Ω for a NAND fault model

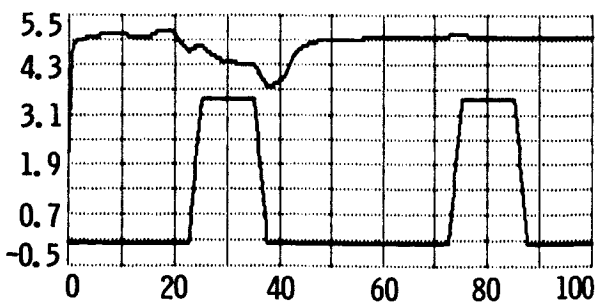


Figure 2.5.2.B-5. R5 = 200 Ω for a NAND fault model

Figure 2.5.2.B-6. R6 = 100 Ω for a NAND fault model

Figure 2.5.2.B-7. R6 = 500 Ω for a NAND fault model

Figure 2.5.2.B-8. R8 = 1,000 Ω for a NAND fault model

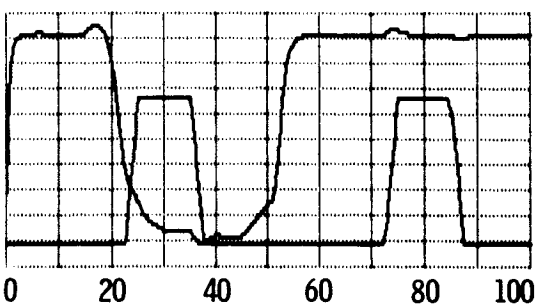


Figure 2.5.2.B-9. R8 = 1,000,000 Ω for a NAND fault model

Figure 2.5.2.B-10. R10 = 1,000 Ω for a NAND fault model

Figure 2.5.2.B-11. R11 = 1,000 Ω for a NAND fault model

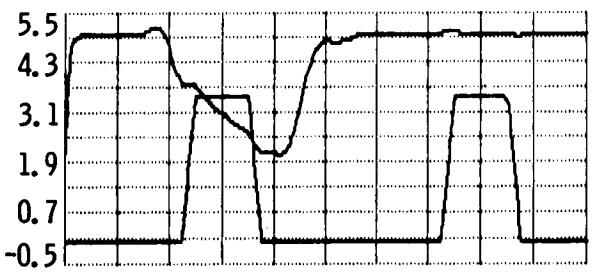


Figure 2.5.2.B-12. R12 = 1,000 Ω for a NAND fault model

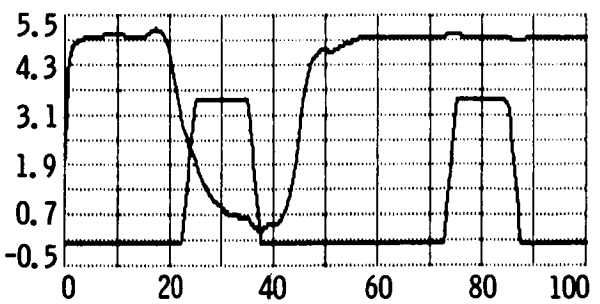


Figure 2.5.2.B-13. R13 = 0.1 Ω for a NAND fault model

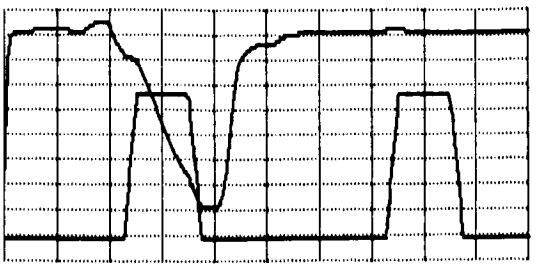


Figure 2.5.2.B-14. R14 = 1,000 Ω for a NAND fault model

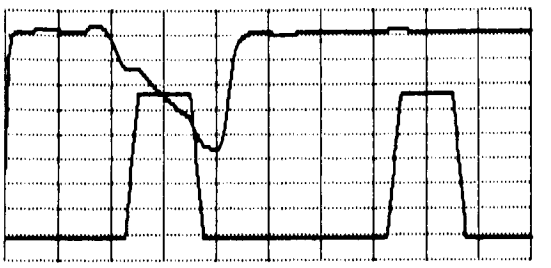


Figure 2.5.2.B-15. R15 = 0.5 Ω for a NAND fault model

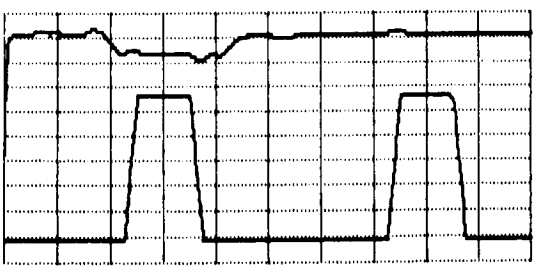


Figure 2.5.2.B-16. R18 = 10 Ω for a NAND fault model

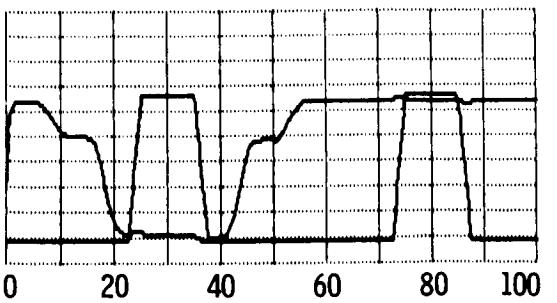


Figure 2.5.2.B-17. R19 = 10 Ω for a NAND fault model

Figure 2.5.2.B-18. R23 = 10 Ω for a NAND fault model

Figure 2.5.2.B-19. R24 = 1,000 Ω for a NAND fault model

Figure 2.5.2.B-20. R24 = 10 Ω for a NAND fault model

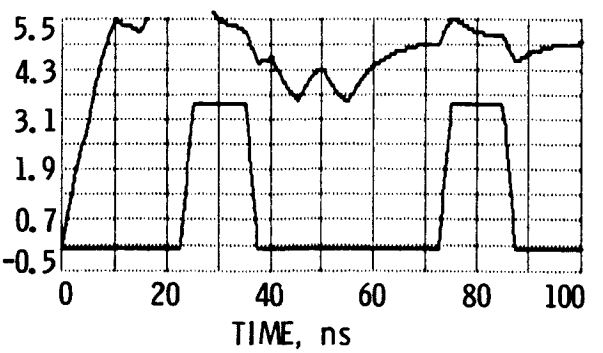


Figure 2.5.2.B-21. R5 = 1,000,000 Ω for a NAND fault model

APPENDIX 2.5.2.C

CMOS NOR TRANSISTOR-LEVEL FAULT MODEL ANALYSIS

The failure effects of a CMOS NOR gate were simulated using a transistor-level analysis program. The behavior of this circuit is analogous to that observed in Appendix 2.5.2.A: stuck-ats, asymmetric timing degradations, signal attenuation, and voltage division. The fault model is illustrated in Figure 2.5.2.C-1. The trace in the figure shows the NOR inputs produced by two of the pulse generators. The circuit contains three pulse generators, four transistors, a coupling capacitor from one of the pulse generators to the NOR output, and 11 resistors and a capacitor defined as:

R1:	Open from input to P-transistor	0.01 Ω
R2:	Open from N-transistor to ground	0.01 Ω
R3:	Open from pulse generator to input	0.01 Ω
R4:	Open from both N-transistors to ground	0.01 Ω
R5:	Open from P-transistor to output	0.01 Ω
R6:	Attenuation from the NOR output	0.01 Ω
R7:	Open from N-transistor to ground	0.01 Ω
R8:	Short from output to ground	1E6 Ω
R9:	Short from Vdd to output	1E6 Ω
R10:	Open from Vdd to NOR gate	0.01 Ω
R11:	Short from input to output	0.01 Ω
C1:	Capacitance	1E-11 F

Figure 2.5.2.C-2 shows the nominal NOR gate behavior. The bottom trace shows a double pulse from the capacitively coupled pulse generator as well as the NOR gate output. The output shows the characteristic quick pull down and slow pull up (due to the series P-transistors).

Resistor R1 was set to 1 kohm in Figure 2.5.2.C-3 to simulate a moderate open from the input to the P-transistor. The NOR gate suffers a slight timing degradation in the 0 to 1 transition.

Figure 2.5.2.C-4 simulates a significant open to a P-transistor. There is a further delay in the 0 to 1 transition of the NOR gate.

Figure 2.5.2.C-5 show the result of a significant open from the leftmost N-transistor to ground. Due the input signal order, the other N-transistor pulls the output low before the faulty transistor turns on. As a result, there is no damage indicated in the NOR gate output.

In Figure 2.5.2.C-6 a moderate open is inserted into the gate lead of the rightmost N-transistor. A slight timing degradation occurs in the 1 to 0 transition.

Figure 2.5.2.C-7 represents the effects of a moderate open in the ground connection. An asymmetric signal attenuation occurs in the 1 to 0 transition.

A moderate open from the P-transistor to the output is shown in Figure 2.5.2.C-8. A severe timing degradation is the primary result.

In Figure 2.5.2.C-9 a moderate open is simulated from the NOR gate output. The coupled signal source dominates the behavior of this circuit.

Figure 2.5.2.C-10 simulates a moderate open in the ground lead of the rightmost N-transistor. This simulation shows a timing degradation in the 1 to 0 transition.

A low impedence short from the output to ground is simulated in Figure 2.5.2.C-11. Here, the NOR gate behaves stuck-at-0.

Figure 2.5.2.C-12 simulates a low impedence short from the output to Vdd. The NOR gate exhibits a stuck-at-1 behavior.

In Figure 2.5.2.C-13 a moderate open was simulated in the Vdd supply path. There is a significant timing degradation in the 0 to 1 transition.

Finally, Figure 2.5.2.C-14 shows the behavior when there is a small impedence from the input to the output. The resulting output from the voltage divider is an attenuated version of the input.

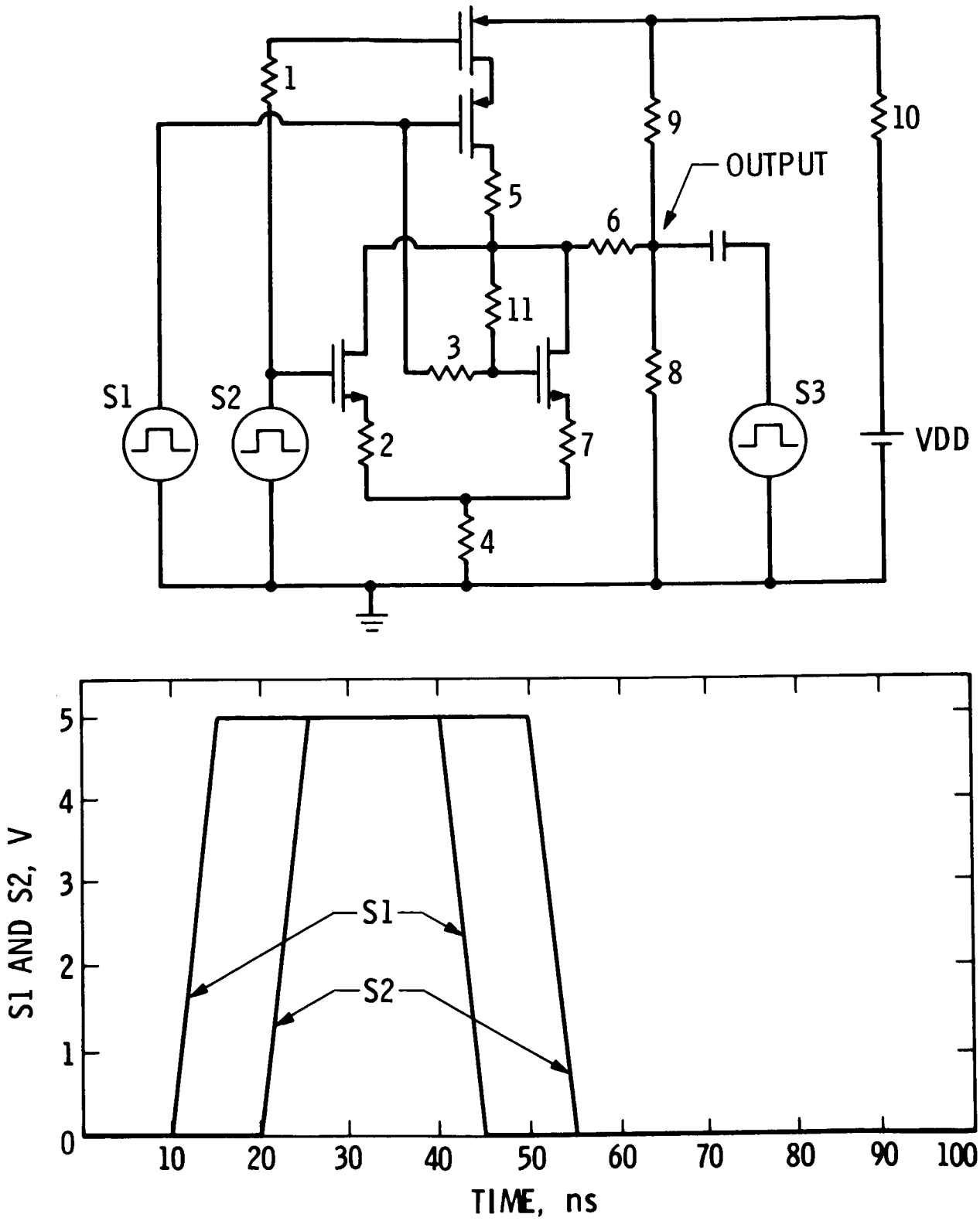


Figure 2.5.2.C-1. NOR fault model

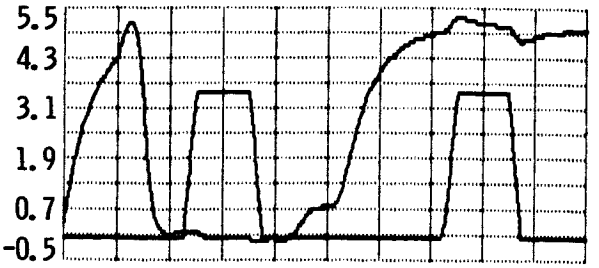


Figure 2.5.2.C-2. Nominal behavior for a NOR fault model

Figure 2.5.2.C-3. R1 = 1,000 Ω for a NOR fault model

Figure 2.5.2.C-4. R1 = 1,000,000 Ω for a NOR fault model

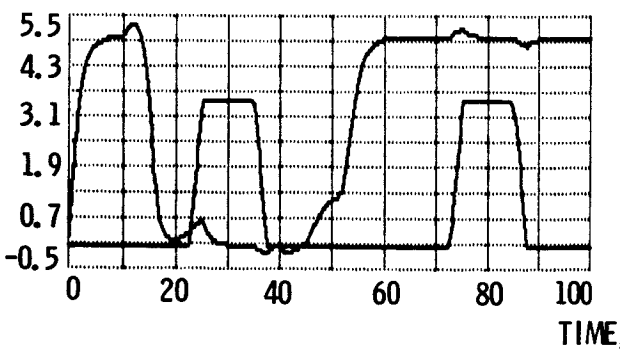


Figure 2.5.2.C-5. R2 = 1,000,000 Ω for a NOR fault model

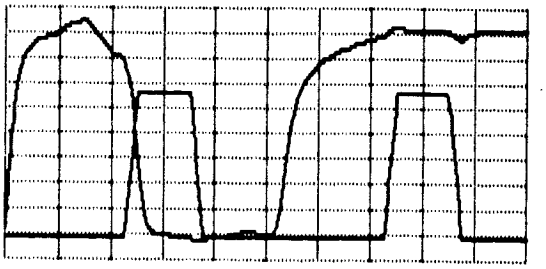


Figure 2.5.2.C-6. R3 = 1,000 Ω for a NOR fault model

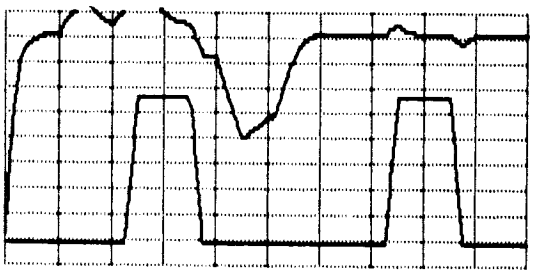


Figure 2.5.2.C-7. R4 = 1,000 Ω for a NOR fault model

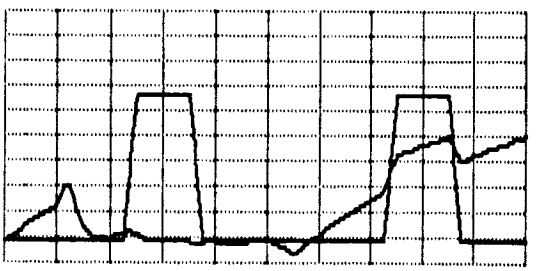


Figure 2.5.2.C-8. R5 = 1,000 Ω for a NOR fault model

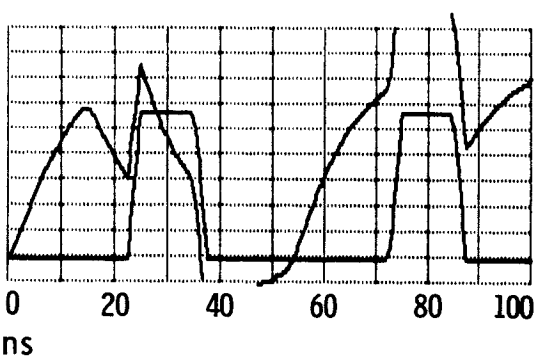


Figure 2.5.2.C-9. R6 = 1,000 Ω for a NOR fault model

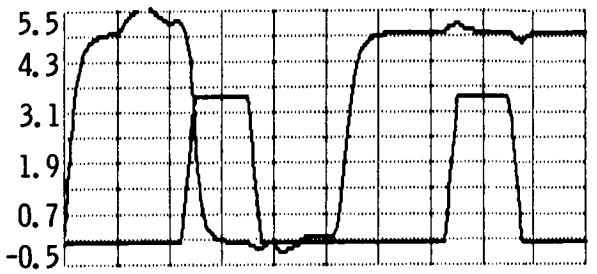


Figure 2.5.2.C-10. R7 = 1,000 Ω for a NOR fault model

Figure 2.5.2.C-13. R10 = 100 Ω for a NOR fault model

Figure 2.5.2.C-11. R8 = 10 Ω for a NOR fault model

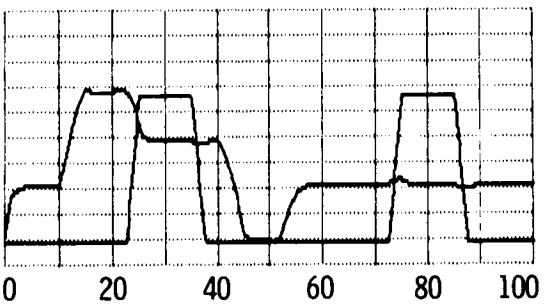


Figure 2.5.2.C-14. R11 = 10 Ω for a NOR fault model

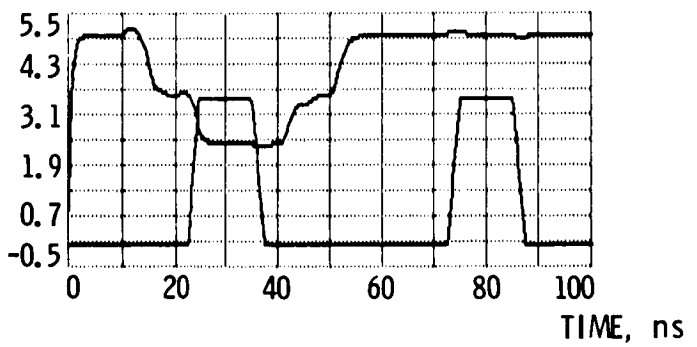


Figure 2.5.2.C-12. R9 = 10 Ω for a NOR fault model

APPENDIX 2.5.2.D

CMOS STATIC FLIP-FLOP TRANSISTOR-LEVEL FAULT MODEL ANALYSIS

The failure effects of a CMOS static flip-flop were analyzed using a transistor-level circuit analysis program. Figure 2.5.2.D-1 shows the circuit model. Nominal component values used in this analysis are listed below. As illustrated in the fault simulations, the flip-flop exhibits stuck-at-1 and stuck-at-0 behavior, asymmetric timing degradations and failures whose effects depend on the switching threshold and timing of external circuits. These latter failures may cause errors under certain operations and may be invisible for other operations. It remains to investigate the effects of these undetectable latent failures on error-checking algorithms.

Figure 2.5.2.D-1 comprises standard cross-coupled inverters and two signal sources, S0 and S1, that simulate true and complement data pulses. The trace shows the outputs from the two independent signal sources. Fourteen resistors are used to simulate opens and shorts and are defined as:

R1:	Open from signal source	50 Ω
R2:	Open to gate of N-transistor	0.01 Ω
R3:	Open to gate of P-transistor	0.01 Ω
R4:	Open between N-transistor and ground	0.01 Ω
R5:	Open between P-transistor and Vdd	0.01 Ω
R6:	Open in cross-coupling	0.01 Ω
R7:	Open in cross-coupling	0.01 Ω
R8:	Short between cross-coupling	1E6 Ω
R9:	Open to N-transistor gate	0.01 Ω
R10:	Open to P-transistor gate	0.01 Ω
R11:	Open between N-transistor and ground	0.01 Ω
R12:	Open between P-transistor and Vdd	0.01 Ω
R13:	Open from signal source	50 Ω
R14:	Open from Vdd	0.01 Ω

Figure 2.5.2.D-2 shows the nominal behavior of the flip-flop.

In Figure 2.5.2.D-3, a moderate (1 kohm) open is inserted into the gate connection of one of the N-transistors. The resulting output briefly rises to logic 1 (at 74 ns) while the data pulse is applied. However, since regeneration does not occur, the flip-flop fails to maintain the 1 state. The output of this circuit appears stuck-at-0 if sampled after the data pulse period, but may be interpreted as a good output (depending on the threshold values of succeeding stages) if sampled near the end of the data pulse.

Figure 2.5.2.D-4 shows the analogous behavior when a P-transistor drive is attenuated by a 1-kohm resistance. Here, the output is stuck-at-1 independent of the data pulses.

A moderate open to ground is simulated in Figure 2.5.2.D-5. The resulting output is similar to that observed in Figure 2.5.2.D-3. Since one of the inverters is faulty, it cannot participate in the regeneration process.

Figure 2.5.2.D-6 shows the analogous behavior when a moderate resistance is inserted into the Vdd path of one of the inverters. Figure 2.5.2.D-6 is very similar to Figure 2.5.2.D-4.

In Figure 2.5.2.D-7, a moderate open has been inserted into the cross-coupling path. The flip-flop has no difficulty in achieving a 0 state but cannot hold a 1 state. The fault may not be observed externally if the output of this circuit is sampled sufficiently quickly after the data pulses.

Figure 2.5.2.D-8 shows the flip-flop behavior when an open is inserted into the other cross-coupling path. Here, the flip-flop cannot hold a 0 state.

In Figure 2.5.2.D-9 a moderate-valued short was inserted between the cross-coupling connections. This short does not affect the flip-flop behavior.

A low-valued short between cross-coupling paths was simulated in Figure 2.5.2.D-10. Here the output stays in the region between 3 and 3.5 volts. Depending on external circuit behavior, this output may appear stuck-at-1, stuck-at-0, or stuck-at-X.

Figure 2.5.2.D-11 shows the results of a moderate open to the gate of the right N-transistor. The flip-flop fails to achieve a 0 state and therefore appears stuck-at-1.

In Figure 2.5.2.D-12, a moderate open was placed into the gate connection of the right P-transistor. The resulting flip-flop output is stuck-at-0.

A moderate open was inserted into the path of the right signal source in Figure 2.5.2.D-13. Here the flip-flop behaves stuck-at-1.

In Figure 2.5.2.D-14, an open was inserted into the Vdd connection of the right inverter. The flip-flop exhibits a stuck-at-0 behavior.

Figure 2.5.2.D-15 shows flip-flop behavior when a 1-kohm series resistor is inserted into the path from the right signal source. There is a slight timing degradation in the 0 to 1 transition.

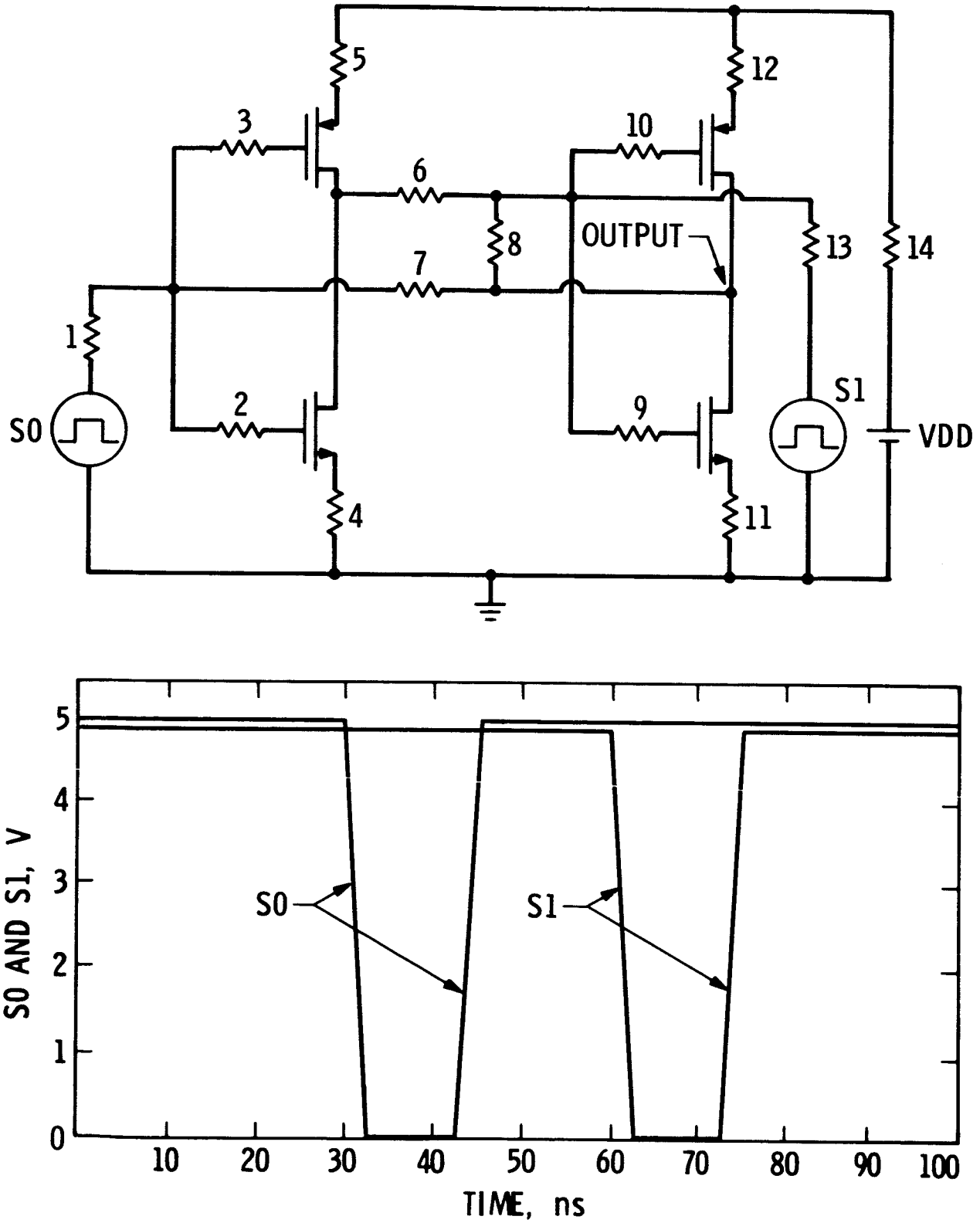


Figure 2.5.2.D-1. Static CMOS flip-flop fault model

Figure 2.5.2.D-2. Nominal behavior for a static CMOS flip-flop fault model

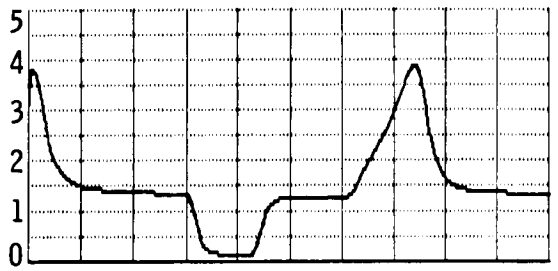


Figure 2.5.2.D-3. R2 = 1,000 Ω for a static CMOS flip-flop fault model

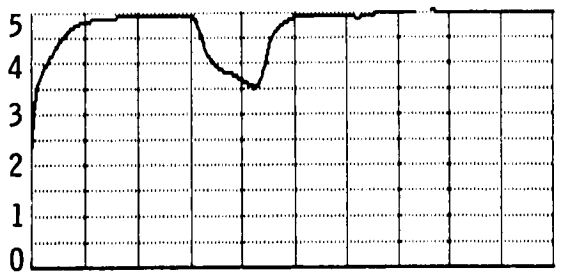


Figure 2.5.2.D-4. R3 = 1,000 Ω for a static CMOS flip-flop fault model

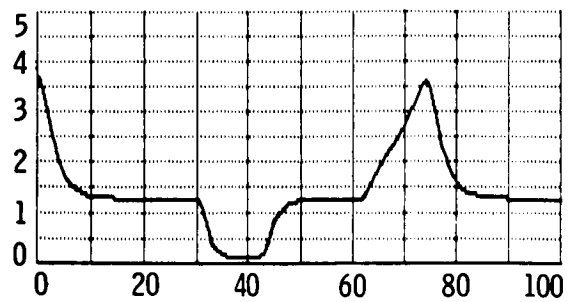


Figure 2.5.2.D-5. R4 = 1,000 Ω for a static CMOS flip-flop fault model

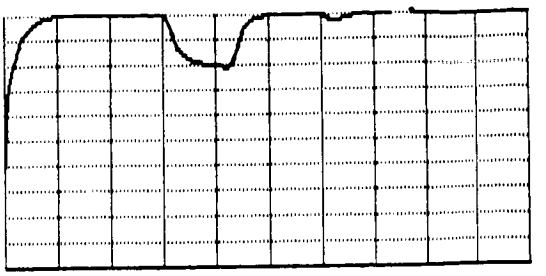


Figure 2.5.2.D-6. R5 = 1,000 Ω for a static CMOS flip-flop fault model

Figure 2.5.2.D-7. R6 = 1,000 Ω for a static CMOS flip-flop fault model

Figure 2.5.2.D-8. R7 = 1,000 Ω for a static CMOS flip-flop fault model

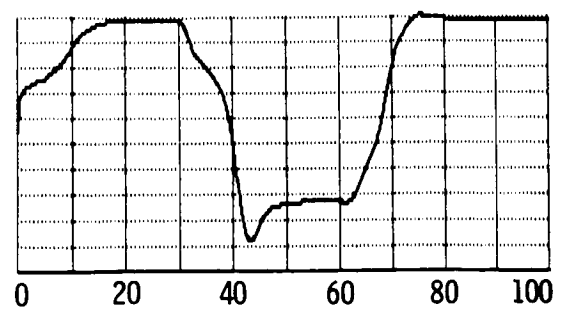


Figure 2.5.2.D-9. R8 = 1,000 Ω for a static CMOS flip-flop fault model

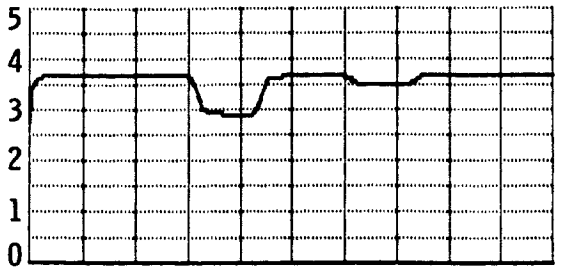


Figure 2.5.2.D-10. R8 = 10 Ω for a static CMOS flip-flop fault model

Figure 2.5.2.D-11. R9 = 1,000 Ω for a static CMOS flip-flop fault model

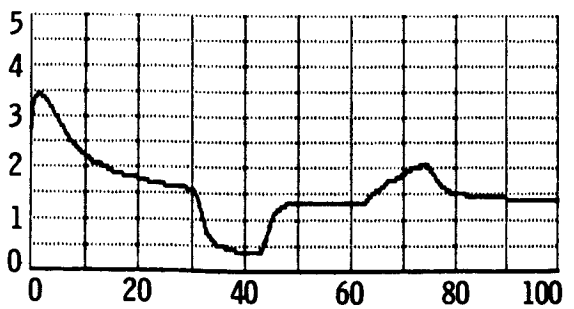


Figure 2.5.2.D-12. R10 = 1,000 Ω for a static CMOS flip-flop fault model

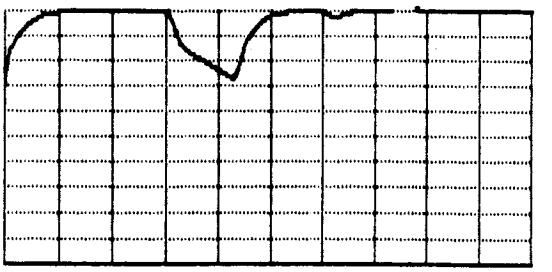


Figure 2.5.2.D-13. R11 = 1,000 Ω for a static CMOS flip-flop fault model

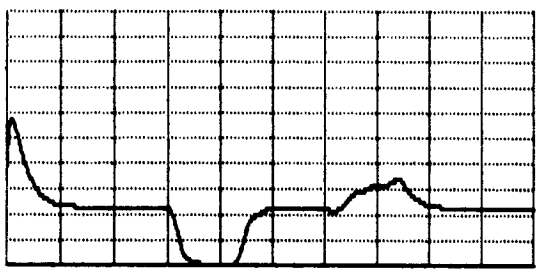


Figure 2.5.2.D-14. R12 = 1,000 Ω for a static CMOS flip-flop fault model

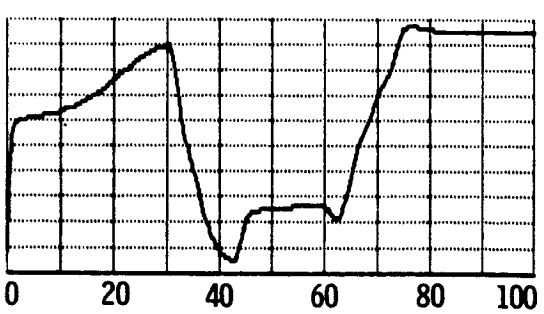


Figure 2.5.2.D-15. R14 = 1,000 Ω for a static CMOS flip-flop fault model

APPENDIX 2.5.2.E

CMOS DYNAMIC LATCH TRANSISTOR-LEVEL FAULT MODEL ANALYSIS

The CMOS dynamic latch element shown in Figure 2.5.2.E-1 was used to simulate various failure mechanisms. The faulty circuit was analyzed using a transistor-level circuit analysis program. Faulty behavior is analogous to that observed in the previous appendixes.

The model in Figure 2.5.2.E-1 comprises an NMOS load transistor configured as a capacitor, a transmission gate, three signal sources, and five resistors. Nominal component values are listed below. The resistors are used to simulate opens and shorts as follows:

R1:	Open to enable on PMOS transistor in transmission gate	0.01 Ω
R2:	Open to enable on NMOS transistor in transmission gate	0.01 Ω
R3:	Short from input to output	0.01 Ω
R4:	Open from transmission gate to capacitor	0.01 Ω
R5:	Short from ground to capacitor	1E7 Ω

Figure 2.5.2.E-2 shows the nominal output behavior of the latch.

In Figure 2.5.2.E-3 the latch behavior is simulated under the condition of a moderate open in the enable to the PMOS transistor in the transmission gate. The p-transistor neither turns on nor turns off fully, leading to a bleeding of the capacitor. The capacitor, in fact, will slowly follow the input data. Hence, this failure looks like a stuck-on-X fault in which X varies between 0 and 1.

The connection to the p-transistor is opened in Figure 2.5.2.E-4. The resulting waveform shows the appearance of both a timing alteration and a signal attenuation.

Figure 2.5.2.E-5 shows the effects of a moderate open to the n-transistor in the transmission gate. Here the n-transistor fails to completely turn on or off, resulting primarily in leakage of the capacitor. This failure is similar to the one shown in Figure 2.5.2.E-3.

Figure 2.5.2.E-6 illustrates the effect of an open to the n-transistor in the transmission gate. The resulting output suffers a timing alteration as well as a signal offset.

A relatively large series resistance from the transmission gate to the capacitor is simulated in Figure 2.5.2.E-7. As is seen, there is virtually no degradation in the latch output.

Figure 2.5.2.E-8 simulates a moderate short to ground. Depending upon the system timing, the latch output might appear stuck-at-zero or, if sampled soon after the latch enable pulse, would appear fault-free.

A moderate short from input to output is shown in Figure 2.5.2.E-9. The dominant characteristic of this failure is a timing degradation.

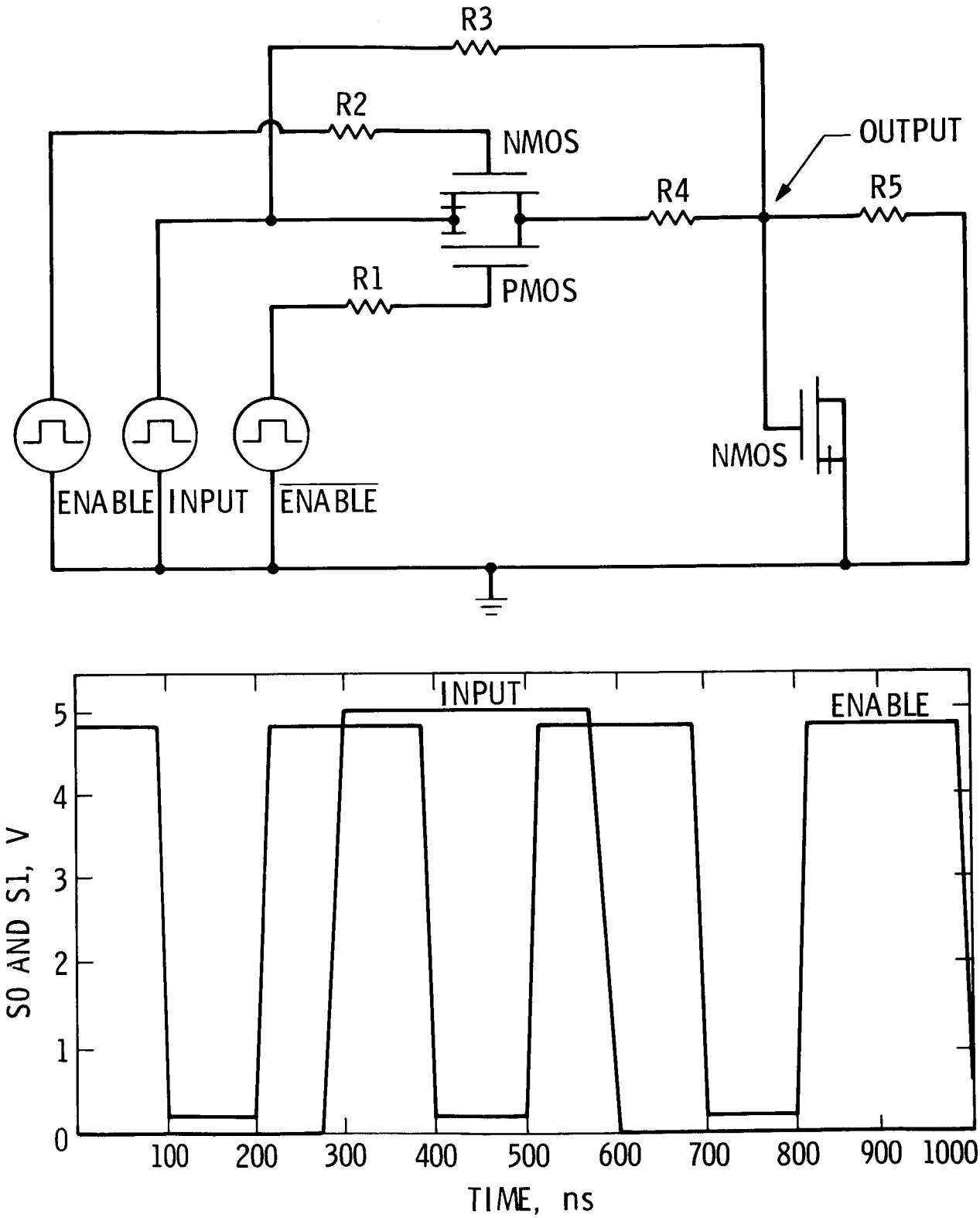


Figure 2.5.2.E-1. Dynamic latch fault model

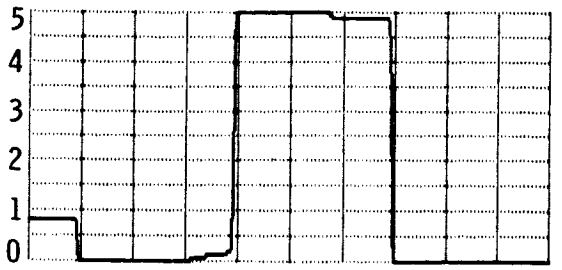


Figure 2.5.2.E-2. Nominal behavior for a dynamic latch fault model

Figure 2.5.2.E-3. R1 = 1,000 Ω for a dynamic latch fault model

Figure 2.5.2.E-4. R1 = 1,000,000 Ω for a dynamic latch fault model

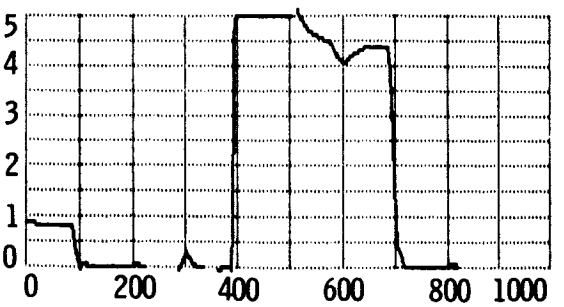


Figure 2.5.2.E-5. R2 = 1,000 Ω for a dynamic latch fault model

Figure 2.5.2.E-6. R2 = 1,000,000 Ω for a dynamic latch fault model

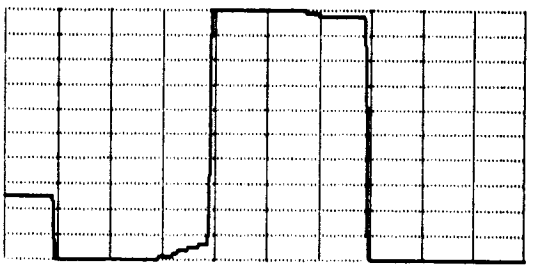


Figure 2.5.2.E-7. R4 = 10,000,000 Ω for a dynamic latch fault model

Figure 2.5.2.E-8. R5 = 1,000 Ω for a dynamic latch fault model

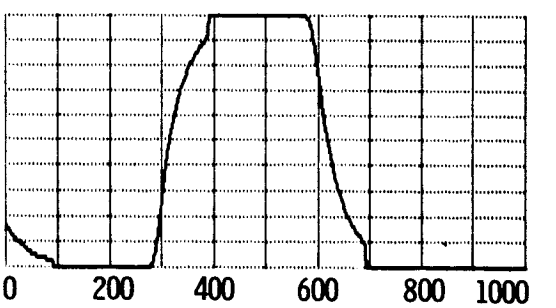


Figure 2.5.2.E-9. R3 = 1,000 Ω for a dynamic latch fault model

2.6 CRRES CHIP[1]

The test circuits described here will be included in the Microelectronics Package (MEP) of the Combined Release and Radiation Effects Satellite (CRRES).

The CRRES chip serves three purposes:

1. It demonstrates the capability of custom LSI/VLSI in a spacecraft flight system environment.

2. It is a vehicle for correlating product assurance technology with device performance.

3. It provides an opportunity to study the effects of single event upset (SEU) and total dose radiation on devices fabricated using foundry processes.

The CRRES chip has three experiments implemented on a single silicon substrate. The circuitry includes a 1-kbit RAM for SEU measurements, a timing sampler for measurement of propagation delay, and an addressable array of transistors for device parameter extraction. The chip also includes test structures to monitor process parameters, and individual transistors identical to those used in the above circuits. These transistors can be probed to characterize the chip before it is packaged.

The CRRES chip will be fabricated using a basic 3-μm oxide isolated p-well CMOS process offered through the MOS Implementation Service (MOSIS) of the USC Information Sciences Institute. Included on the wafers will be "drop-in" test structures for assessment of wafer yield and reliability. Identical wafers will also be fabricated using a "field-hard" process offered by a commercial vendor. This process will yield devices functional up to a total radiation dose of about 500 krad Si.

The chips will be delivered from the fabricators in wafer form. The CRRES chip and the test structures on the wafers will be probed at JPL and subsequently diced and packaged by the JPL packaging group. About 120 packaged chips will be subjected to qualification screening and burn-in to select the flight parts.

Figure 2.6-1 shows the 1k and 4k RAM versions of the CRRES chip floor plan with pin functions labeled. Figure 2.6-2 shows the layout of the CRRES chip and associated test chips for obtaining reliability (time-dependent dielectric breakdown), yield, and parametric data. This composite structure will be repeated across the wafer and will allow a comparison of test chip results with CRRES chip performance and yield. Also included are four inverter circuits configured in a CD 4007 arrangement. These circuits, designed for the Goddard Space Flight Center, will also be included in the CRRES MEP.

[1]This section was prepared by R. H. Nixon.

The CRRES chip is packaged in a 64-pin dual in-line package. The table in Appendix 2.6.A describes the pin functions in detail and indicates the type of interface to the spacecraft system.

2.6.1 Random Access Memory (RAM)

The CRRES chip static RAM has been designed specifically for the detection of single event upsets. The RAM architecture is based on the conventional static-CMOS six-transistor RAM cell found in many commercial designs. The sophisticated timing circuits for enhancing the speed of these commercial designs has been removed to improve circuit reliability in a space radiation environment. Both a 1-kbit and a 4-kbit version of the RAM will be built.

2.6.1.1 RAM Operation. The RAM is organized as 64 or 256 16-bit words. It has a 16-bit bidirectional data bus for reading and writing, a 6- or 8-bit address bus and four input control lines. The only functional difference between the 1-kbit and the 4-kbit RAM is the use of two additional address lines. These lines are not used on the 1-kbit RAM, but may be connected to active circuitry if desired without harming the RAM function.

The data bus is connected to chip pins 30 to 45, with pin 45 being the least significant bit. The address bus is connected to pins 50 to 55, with pin 55 being the least significant bit. The 4-kbit version of the RAM has two additional address bits: A6 at pin 57 and A7 at pin 3.

The RAM schematic (Figure 2.6-3) and the RAM timing diagram (Figure 2.6-4) illustrate the function and operation of the RAM. RAM pin connections are shown in Figure 2.6-3 as well as in the table in Appendix 2.6.A.

The S control line (pin 46) acts as a chip select for the RAM only. The other circuits on the chip are not affected by S. When S is low, the data bus is driven to a high impedance mode, and the address bus and remaining control lines have no effect on the RAM. When S is high, the W, E, and EP control lines are enabled.

The W control line (pin 47) is the read/write control line. When W is low, the RAM is in the read mode and the tristate I/O is at low impedance and drives DQ according to the output of the RAM column data latch. When W is high, the RAM is in a write mode and the tristate I/O pad is in the high impedance state. Data at DQ is connected through the transmission gates and into the memory cell.

The E control line (pin 48) enables the memory cell row select and read/write circuits when it is high. When the RAM is in read mode, the column data latches follow the bit lines when E is high and hold data when E is low. When the RAM is in the write mode, the write data transmission gates are on when E is high and open when E is low.

The EP control line (pin 49) controls the column bit line precharge. Since the memory cells are coupled to the bit lines with n-channel transistors, the bit lines must be precharged to a logical 1 prior to reading data. EP has no effect if the RAM is not selected (S low). In this case, the precharge transistors are always on. When EP is low and the RAM is selected, the precharge transistors will remain on until E goes high.

The RAM shares Vdd with the chip Vdd (pin 1) and Vss with the chip Vss (pin 29). From tests on prototype circuits, we have determined that the RAM will operate at clock periods as small as 400 ns using the timing sequences shown in Figure 2.6-4. Prototype RAM circuit test results are:

IDD standby:	< 1 μA
IDD selected:	5 μA
IDD operating: (2-μs clock rate)	3.5 μA
Read/Write access time	70-80 ns
Chip select time	200 ns

2.6.1.2 RAM SEU Analysis. The RAM memory cell has been designed with transistor source drain areas that are the minimum allowable for reliable circuit operation under the design rule constraints for the 3-μm CMOS process. However, to facilitate SEU modeling, four different memory cell geometries were designed and were included on the prototype parts used for ground testing. Only the minimum geometry design is included on the flight parts. Within these designs the feature sizes of both the n-channel and p-channel transistors are scaled from minimum geometry to roughly three times the minimum scale size, or about 10 times the area for the source and drain diffusion/ion implantation. This approach provides maximum SEU sensitivity (minimum geometry cells) and enough size variation to observe the effects of scaling on SEU sensitivity. The memory cell transistor designs are also included in the test strip area of the CRRES chip. This will allow direct measurement of device parameters required for SEU modeling.

RAM cell SEU is presently being modeled to predict the critical charge for SEU using SPICE 2G.1. Figure 2.6-5 shows the schematic diagram and the geometrical layout of the mimimum geometry RAM cell. In Figure 2.6-5, the sensitive drain junctions, corresponding to one state of the memory cell, are highlighted as shaded areas. The injection of charge due to a heavy ion striking this sensitive area is modeled as a current source, as shown in the diagram.

Predictions of the SEU rate for the minimum geometry cell have been made based on the Peterson Model [1, 2]. These results are shown for varying values of Vdd in Figure 2.6-6. The critical charge (Qc) required to upset the minimum geometry cell is approximately 0.5 to 0.8 pc, depending on Vdd. By estimating the density of incident radiation with energy above Qc expected in the satellite orbit, the upset rate is estimated to be approximately 0.3 upsets per month per 1024 RAM cells. Prototype CRRES RAM chips are being tested with radiation of known energy and density to confirm the model predictions.

2.6.2 Transistor Matrix

The transistor matrix allows a statistically significant number of n- and p-channel transistors of varying geometries to be characterized with minimal system overhead and external chip pins.

The matrix, shown in Figure 2.6-7, consists of 16 p-channel and 14 n-channel MOS transistors, of various geometries, arranged in a 4 X 8 array. Two of the matrix locations contain field oxide transistors. Appropriate selection circuitry is included to permit measurement of individual transistor characteristics. The row and column selection circuitry is the same design as that used in the RAM, i.e., NOR logic with a weak n-channel pull down transistor used as a common resistive load. The transistor geometries are illustrated in Figure 2.6-8. These geometries are identical for both n and p types of transistors.

The organization of a small section of the transistor matrix is illustrated in Figure 2.6-9. In this figure, two adjacent columns of a single row are shown in the manner in which they are physically laid out on the chip. A column contains transistors of only one type, i.e., n-channel or p-channel. The n- and p-columns are alternated, with each column containing four identical transistors of one of the types shown in Figure 2.6-7. This placement of transistors is also illustrated in Figure 2.6-9.

The transistor matrix requires eight external pin connections excluding the chip Vdd and Vss. It is enabled by the XT Enable line (pin 56). When the XT Enable line is low, the transistors are biased to a known state (p-channel transistors turned off and n-channel transistors turned on). These states were chosen because they are the worst-case bias conditions that will cause the greatest shift in transistor threshold voltage. When the XT Enable line is high, the matrix is enabled, allowing measurements on individual transistors to be made. The XT Enable line enables only the transistor matrix and has no effect on the other circuits on the chip.

Selection of an individual transistor is achieved by addressing the row and column of the transistor. The R0 and R1 lines (pins 59 and 58, respectively) select the rows 0 through 3, with R0 being the least significant bit. The C0, C1, and C2 lines (pins 60, 61, and 62, respectively) select the columns 0 through 7, with C0 being the least significant bit. All row and column select lines are true with a logical 1.

Transistor measurements are made by selecting the matrix, addressing a transistor (pins 58 to 62), supplying a current to the ID line (pin 63) and a voltage to the VG line (pin 4), and measuring the voltage on the VD line (pin 64). If the current to the ID line is measured and the voltage on the VG line is known, the transistor drain voltage, drain current, and gate voltage can be determined directly. By varying the voltage on the ID and VG lines, different operating points can be characterized.

The transistor matrix has three lines available for evaluation purposes that will be connected to Vdd or Vss during actual operation. In the CRRES microelectronics package (MEP), the VPS line (pin 2) is connected to +5 volts and the XT Well (pin 11) line is connected to ground (Vss). For ground testing, these pins will be biased so that a back gate bias is placed on the transistors to measure the body effect parameter. The transistor matrix select circuit (all transistors excluding the transistors to be tested) is powered by the chip Vdd (pin 1).

Transistor parameters will be extracted from a matrix of drain current, gate voltage, and drain voltage measurements from individual transistors. The resulting matrix of transistor parameters for different-sized devices will then be solved for the scale-dependent global transistor parameters (see Section 2.3.6). The transistor measurement configuration is shown in Figure 2.6-10. This is a schematic representation of the layout shown in Figure 2.6-7. The matrix was designed so that the drain voltage, noted as VD(measure), is measured using a Kelvin connection. This can be seen by a close inspection of Figure 2.6-7 where VD(measure) taps the voltage very close to the point where the drain current ID enters the transistor. The overall leakage current for the matrix is less than 10 pA, which is quite acceptable for transistor characterization. It can be see from Figure 2.6-11 that the transistors are drawing hundreds of microamperes of current. Figure 2.6-11 shows data taken from a prototype fabrication run.

It remains to be seen how this matrix will behave after large doses of radiation. Two effects will dominate the behavior with radiation: The transistor threshold voltages will shift and the leakage currents will increase. The shift in the field-oxide transistor thresholds will cause the formation of inversion layers and loss of transistor isolation. The shift in gate-oxide transistor thresholds will result in loss of control in the addressing circuitry. Finally, the increase in bulk leakage with radiation will also result in loss of control in the addressing circuitry.

2.6.3 Timing Sampler and Ring Oscillator

Two types of asynchronous circuits have been included on the CRRES chip for determining the delays of logic transitions through inverters. These circuits are the ring oscillator and the timing sampler. Only the timing sampler will be measured on the satellite; the ring oscillator is included for comparison of ground-based measurements.

The ring oscillator is a commonly used circuit for determining the propagation delay of logic gates. The ring oscillator, however, has been shown to sometimes give erroneous results. In particular, there is the possibility of higher harmonic oscillation modes that, if incorrectly interpreted, can lead to the false calculation of gate delay time [3, 4]. The timing sampler (Figure 2.6-12), which consists of a chain of 128 inverters, is free of this problem because delays are directly measured by means of externally generated timing events (transitions) as opposed to internally generated timing events.

2.6.3.1 Circuit Operation. The ring oscillator consists of a ring of 257 inverters. It has no stable condition and will oscillate with a period that is some odd submultiple of the delay time twice around the ring. The ring oscillator is tested by initializing it into its fundamental mode of oscillation and then, while assuming that it stays in this mode, measuring the frequency with a frequency counter.

The timing sampler shown in Figure 2.6-12 consists of an inverter chain in which the output of inverter pairs is connected to latches, which in turn are connected to a 6-bit decoder. The timing sampler is tested by applying a start transition to the input, waiting a known delay (tp or tn),

applying a stop transition to the input, and then sampling the binary output of the decoder. This provides for a fast, repeatable, and all-digital measurement of gate delay.

The circuits that detect the output of the inverters consist of a latch circuit followed by a mutual exclusion circuit (Figure 2.6-13). Each latch has two inputs (a and b), one connected to a tap on the delay line and the other to a common enable input. The latches, which are initially reset to the same state, are successively tripped as a signal transition (originating from the transition on the start input) travels down the inverter chain and reaches each tap. All latches, however, can be disabled from tripping at any time after the start transition by means of a stop transition on the enable input. As a result, the first latch through the Nth latch will be set, while the other latches remain reset. The latch outputs are decoded to the binary value N, which is sampled at the circuit probe pads. The delay per stage is then calculated by dividing the total delay (the time elapsed from the start transition to the stop transition) by N. Since some of the tap nodes may be in the process of switching when the latches are disabled, the resulting non-digital voltages being sampled at this time will show up at the latch outputs. One of these latches could possibly be in a metastable state, so that this latch output would remain non-digital indefinitely. To prevent such non-digital behavior from getting into the decoder and showing up in the binary output, the output of each latch is passed through a mutual exclusion circuit.

The mutual exclusion circuit output will not begin to change state until the latch is off balanced from its metastable point by at least one transistor threshold (approximately 1 volt). These circuits also sharpen the edges of the latch outputs that are switching, thus reducing the probability of sampling the binary output while a bit is changing. Since the timing sampler is essentially a delay-to-digital converter, there is a possible quantization error in the measured output N. This error is minus one least significant bit and results in a worst-case stage delay error of 100/N percent.

Inverter-pairs were used as delay elements (stages) in the timing sampler as well as the ring oscillator. This allowed the measured delays obtained from the ring oscillator to be directly compared with those obtained from the timing sampler. The first inverter in the pair drives an identical inverter and thus has a fanout of one, whereas the second inverter drives two inverters and has a fanout of two.

The operation of the timing sampler is controlled by the TSE (pin 16) enable and TSI (pin 15) inverter input lines. As seen in the timing diagram shown in Figure 2.6-13, the timing sampler is in a reset state when TSE and TSI are both high or both low. If TSE changes state, the latches are armed. When TSI changes state, the logic signal corresponding to the new state of TSI ripples through the string of inverter pairs. When TSE changes state again, the latches are disabled and the outputs D0 through D5 show the number of stages the signal rippled between the state changes of TSI and TSE.

2.6.3.2 Circuit Analysis. An analysis of propagation delay and transistor transit time is given below. For the purpose of this analysis, the following naming conventions have been used:

TPD = propagation delay time
TAU = transistor transit time
ip = inverter pair
\- = falling edge
\+ = rising edge
_ = underscore--this is used in conjunction with other designators to characterize the type of delay time or transit time that is being described, e.g., TPDipt is the propagation delay time for the rising edge of an inverter pair.

The circuit tau model [5] essentially says that the propagation delay of a negative or positive logic transition through an inverter is proportional to the ratio of the inverter load capacitance to the gate capacitance of the active transistor, i.e., the n-channel transistor gate capacitance for the inverter output's falling edges, or the inverter p-channel transistor gate capacitance for the inverter output's rising edges. The proportionality constants for these delays are the tau of the n-channel transistor (TAUn) and the tau of the p-channel transistor (TAUp), respectively. The TAUn and TAUp can be determined for the timing sampler or ring oscillator by analyzing their response.

2.6.3.2.1 Loaded Inverter Pair. Consider the loaded inverter pair shown in Figure 2.6-14. All transistors are of the minimum length (Ln = Lp = minimum) allowed by the design rules. The loading of the inverter can be expressed as the product of the input gate capacitance of an inverter (Cn + Cp) and the fanout. The fanout of the first stage of an inverter pair (node "a") is 1, and that of the second stage (node "b") is f. The ratio r of an inverter is defined as the width-to-length ratio of the p-channel transistor to that of the n-channel transistor, or r = (W/L)p/(W/L)n. Since Lp = Ln, r is also equal to the ratio of the gate capacitance of the p-channel transistor to that of the n-channel transistor (r = Cp/Cn). The values of f and r in the CRRES timing sampler and ring oscillator designs are f = 2, and r = 5/3. The delays through the first and second inverter can now be expressed in terms of TAUn, TAUp, r, and f. The propagation delays through the first inverter for the falling and rising edges at node "a" are, respectively:

$$TPDa- = \frac{Cn + CP}{Cn} TAUn = (1 + r)\ TAUn$$

and

$$TPDa+ = \frac{Cn + CP}{Cp} TAUp = (1 + 1/r)\ TAUp.$$

The delay through the second inverter for the falling and rising edges at node "b" are, respectively:

$$TPDb- = \frac{f(Cn + Cp)}{Cn} TAUn = f(1 + r)\ TAUn$$

and

$$TPDb+ = \frac{f(Cn + Cp)}{Cp} TAUp = f(1 + 1/r)\ TAUp.$$

By summing the individual inverter delays, the inverter-pair propagation delays for negative and positive edges at node "b" are, respectively:

$$\begin{aligned} TPDip- &= TPDa+ + TPDb- \\ &= (1 + 1/r)\ TAUp + f(1 + r)\ TAUn \end{aligned} \quad (1)$$

and

$$\begin{aligned} TPDip+ &= TPDa- + TPDb+ \\ &= (1 + r)\ TAUn + f(1 + 1/r)\ TAUp. \end{aligned} \quad (2)$$

2.6.3.2.2 Timing Sampler. The inverter pair delays determined from the timing sampler measurements are:

$$TPDip- = tn/N-$$

$$TPDip+ = tp/N+$$

where tn and tp are the applied delays (see Figure 2.6-12) and N- and N+ are the number of latches tripped for negative and positive transitions, respectively. Solving Equations 1 and 2 for the two unknowns, TAUn and TAUp, yields:

$$TAUn = \frac{f \cdot TPDip- - TPDip+}{(1 + r)(f^2 - 1)} \qquad (f > 1) \quad (3)$$

and

$$TAUp = \frac{f \cdot TPDip+ - TPDip}{(1 + 1/r)(f^2 - 1)} \qquad (f > 1) \quad (4)$$

2.6.3.2.3 Ring Oscillator. The ring oscillator consists of loaded inverter pairs and a NOR gate for initialization (see Figure 2.6-15). The on time and off time of the oscillation waveform are:

$$t(on) = n \cdot TPDip+ + (1 + r)\ TAUn$$

and

$$t(off) = n \cdot TPDip- + (1 + 1/r)\ TAUp$$

where n is the number of inverter pairs in the ring oscillator.

Applying Equations (1) and (2) yields:

$$t(on) = n[(1 + r)\ TAUn + f(1 + 1/r)\ TAUp] + (1 + r)\ TAUn \quad (5)$$

$$t(off) = n[f(1 + r)\ TAUn + (1 + 1/r)\ TAUp] + (1 + 1/r)\ TAUp \quad (6)$$

The number of inversions in the ring oscillator is 2n+1, which includes the n inverter pairs plus the NOR gate initialization stage.

The period of oscillation T is: $T = t(on) + t(off)$.

The solution to Equations 5 and 6 for the two unknowns, TAUn and TAUp, yields:

$$TAUn = \frac{n \cdot f \cdot t(off) - (n + 1) \cdot t(on)}{(1 + r)[n^2 \cdot f^2 - (n + 1)^2]} \quad (7)$$

$$TAUp = \frac{n \cdot f \cdot t(off) - (n + 1) \cdot t(on)}{(1 + 1/r)[n^2 \cdot f^2 - (n + 1)^2]} \quad (8)$$

Table 2.6-1 shows test results taken from test circuits fabricated in CMOS. Delays were measured for both positive and negative transitions. The difference between these delays is a function of the p-channel and n-channel transistor channel mobility difference, the ratio of the p-channel to n-channel width to length ratio (r), and the fanout of the inverter pairs (f). These delays turned out to be equal for this CMOS/bulk design ($r = 5/3$, $f = 2$). Values for the n-channel transistor tau (TAUn) and the p-channel transistor tau (TAUp) are calculated from these delays.

2.6.3.2.4 Conclusions. Measurements taken on the timing sampler and ring oscillator show good correlation. The results demonstrate that the timing sampler approach can yield good values for propagation delay and tau. Some minor inadequacies should be noted, however. The timing sampler design used on the CRRES does not take into account parasitic capacitances, and suffers from a worst-case quantization error of plus or minus one bit. Future designs are being planned that will eliminate the error and take into account the parasitic capacitances.

2.6.4 Chip Fabrication History

To date, several prototype designs have been fabricated and tested. In some of the designs, circuits were non-functional because of design errors. In all the prototype fabrication runs, circuits on some chips were non-functional because of mask or fabrication problems. A summary of the fabrication of devices built so far appears in Table 2.6-2.

2.6.5 Chip Requirements and Plans for CRRES

A total of 36 packaged chips will be delivered to CRRES in June of 1985. Twelve chips will actually fly on board the spacecraft. The remaining 24 will be used for ground testing and spares. The plan calls for four of the flight chips to come from the radiation-hard process and the remaining chips from the MOSIS foundries. All chips will be packaged and tested at JPL. Some prototype chips and spares will undergo limited total dose and SEU testing to further verify the design and data reduction procedures.

A brief description of the testing strategy for the three major functions is given below.

RAM: The CRRES RAM is one of many other RAMs that will be sampled periodically by the Microelectronics Package to determine if single event upsets have occurred. A known pattern will be written into the RAM and subsequently interrogated to determine if a bit (or cell) has flipped. Any difference between the pattern written and the pattern read will be transmitted to ground stations for anaylsis.

XT MATRIX: Each of the 30 transistors in the matrix will be measured in the configuration shown in Figure 2.6-10. A total of 120 data points per transistor will be measured. ID will be measured for 10 settings of VD and 12 settings of VG. The matrix will be sampled according to the satellite sampling plan. The current ID will be resolved to $\pm$ 1 μa, and the VD(measure) to $\pm$ 10 mV. The voltages VG, VD(force), and Vdd will be held to within $\pm$ 1 mV.

TIMING SAMPLER: The timing sampler requires a start and stop signal to determine the precise sampling period for resolving the propagation delay. This period will be approximately 100 ns $\pm$ 1 ns. The number of stages that are tripped will be represented by a 6-bit binary word. The timing sampler will be interrogated according to the satellite sampling plan.

2.6.6 References

1. P. Shapiro, E. L. Peterson, and J. H. Adams, Jr., "Calculation of Cosmic-Ray Induced Soft Upsets and Scaling in VLSI Devices," NRL Memo Report No. 4864 (August 26, 1982).

2. J. C. Pickel and J. T. Blandford, Jr., "Single Event Upset Modeling for Static MOS Memory Cells," Rockwell International Report SC5308.9FRD (February 1983).

3. N. Sasaki, "Higher Harmonic Generation in CMOS/SOS Ring Oscillators," IEEE Trans. Electron Devices, ED-29, 280-283 (1982).

4. T. W. Houston, "Comments on Higher Harmonic Generation in CMOS/SOS Ring Oscillators," IEEE Trans. Electron Devices, ED-30, 958-959 (1983).

5. C. L. Seitz, Class notes from Caltech VLSI design class cs181 (see Appendix 2.6.B).

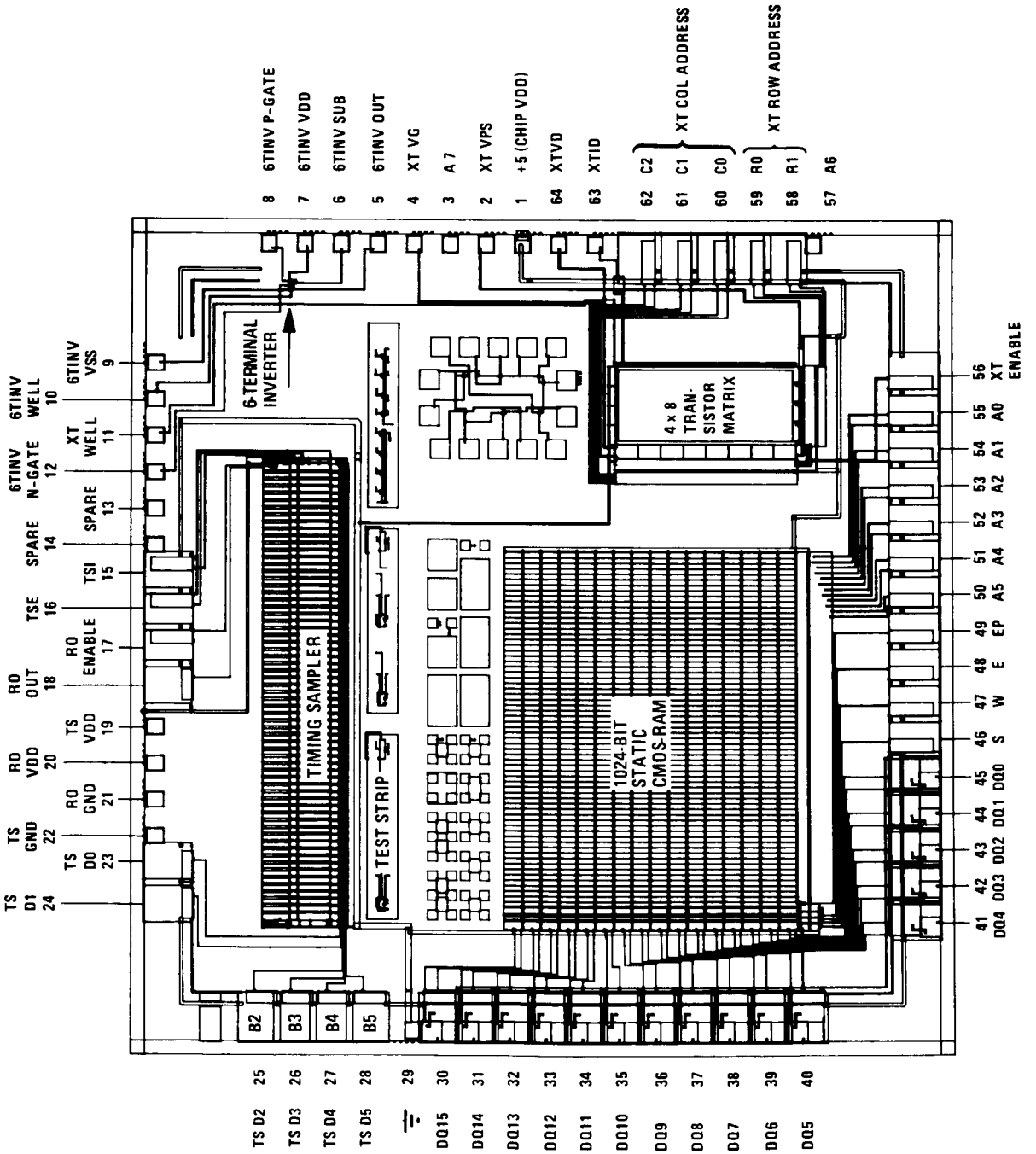


Figure 2.6-1a. CRRES test chip 1k RAM floor plan

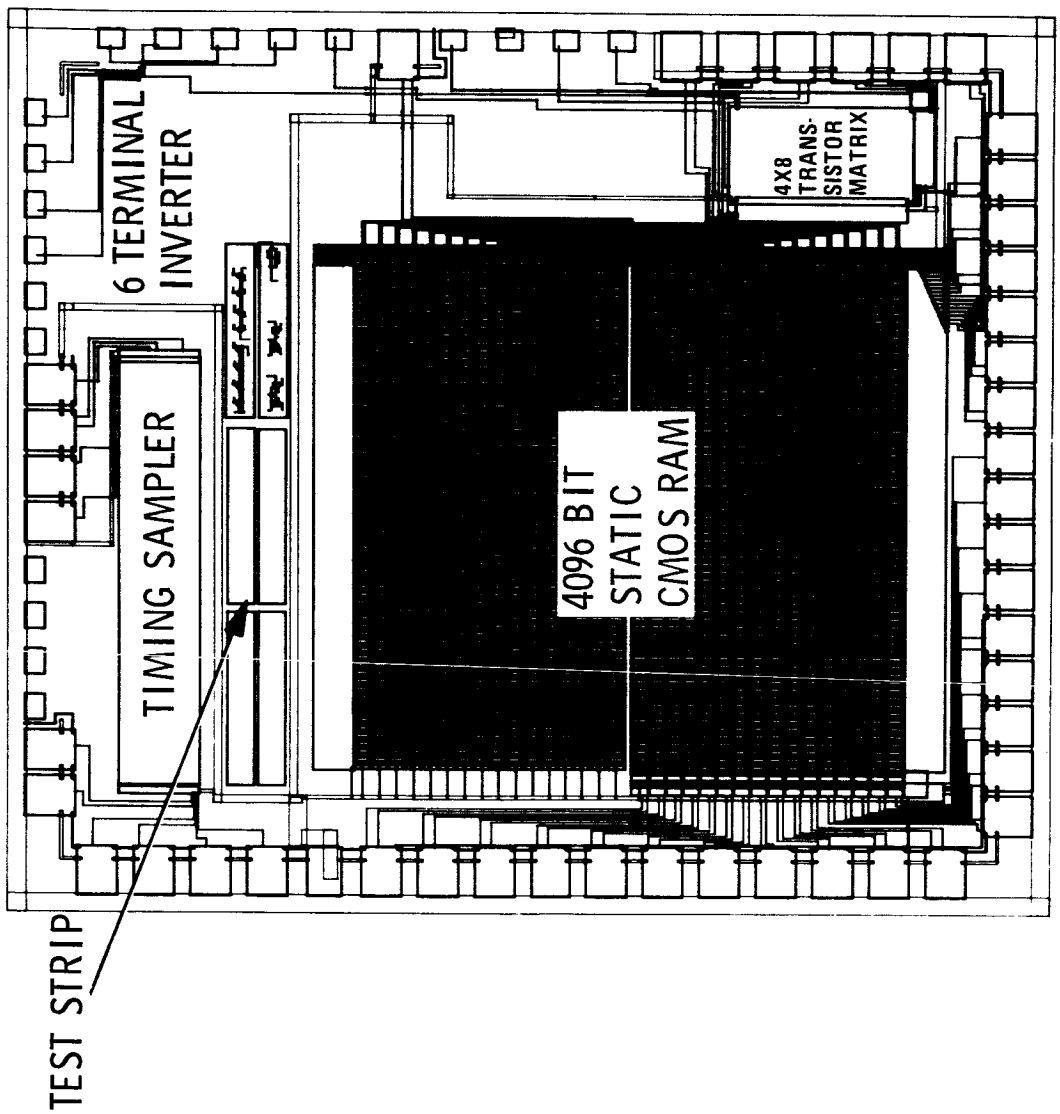


Figure 2.6-1b. CRRES test chip 4k RAM floor plan

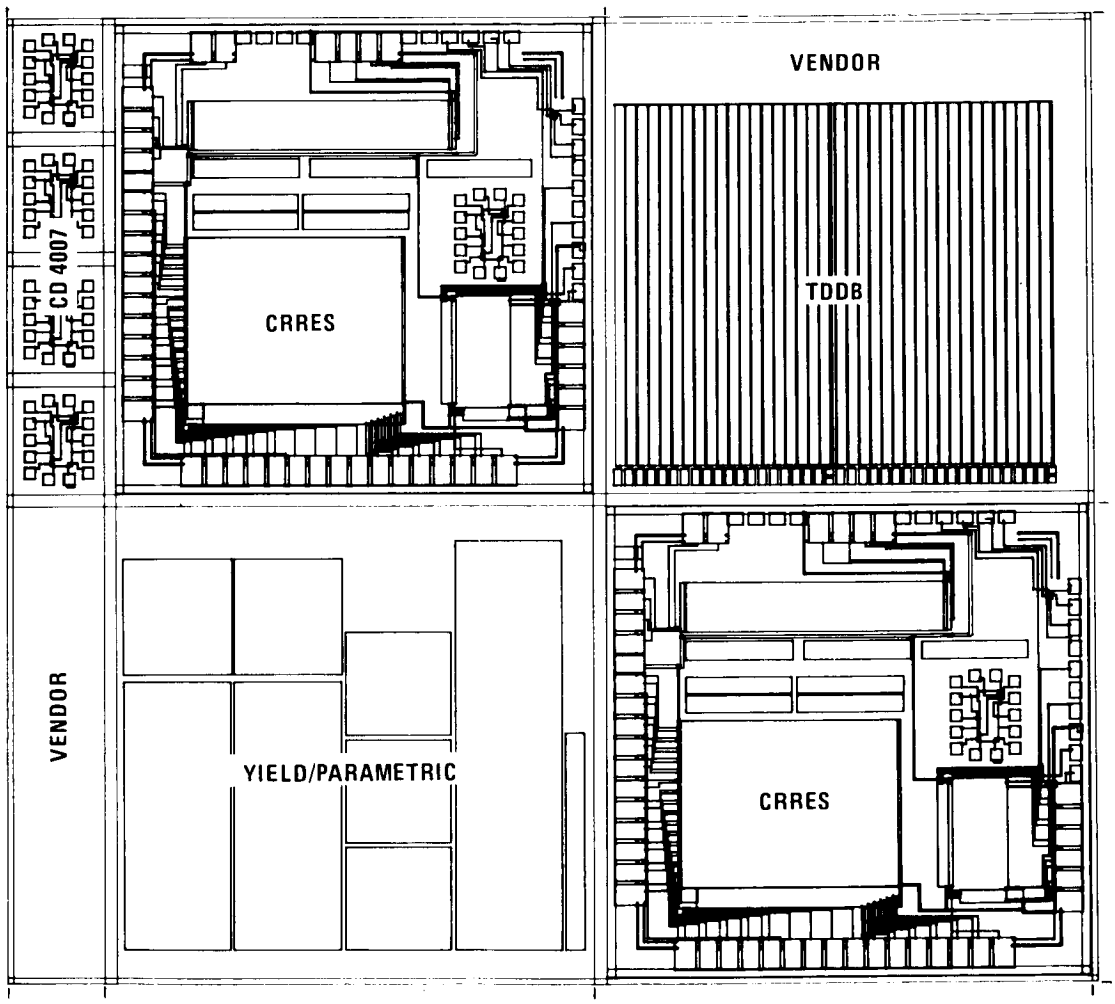


Figure 2.6-2. 3-μm CMOS-bulk chip set

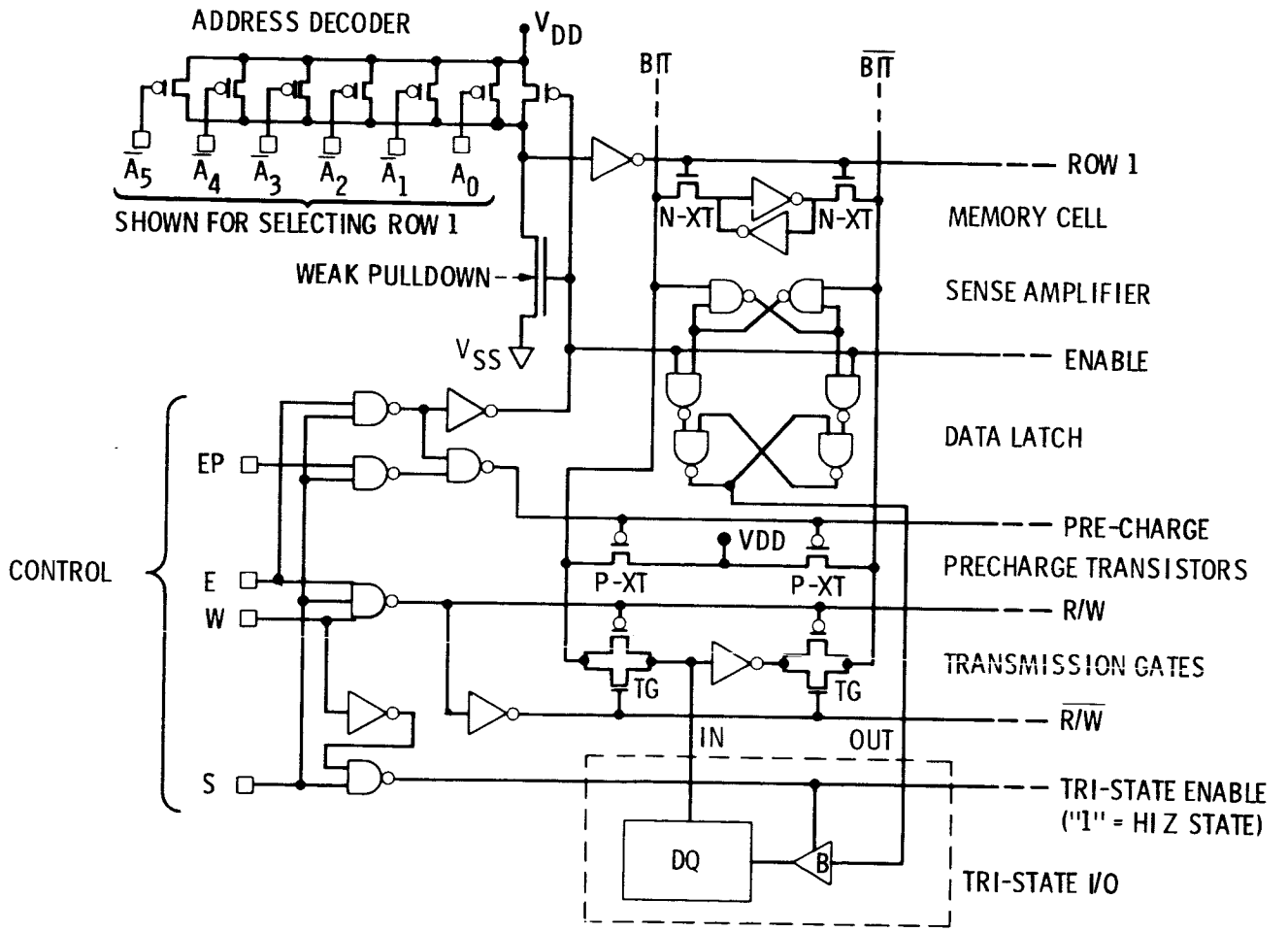
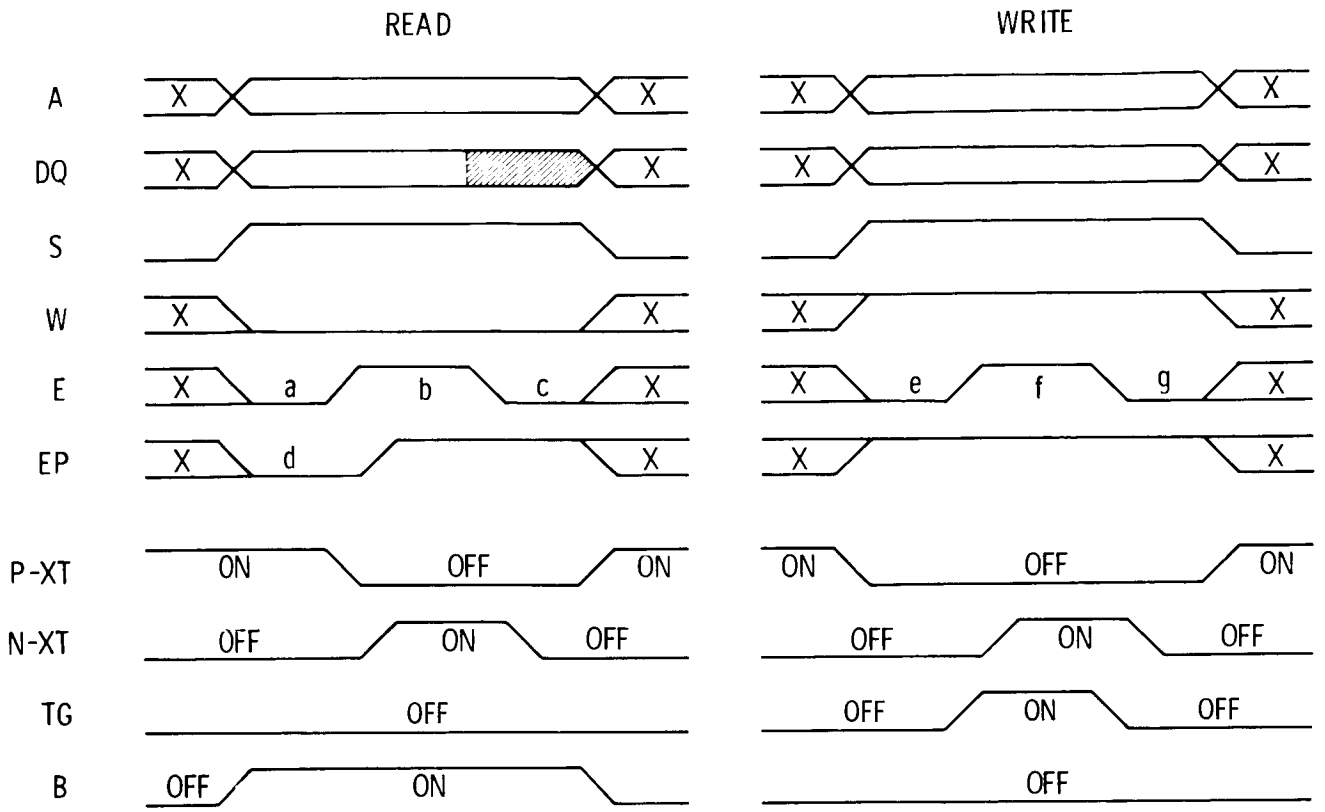


Figure 2.6-3. RAM logic diagram

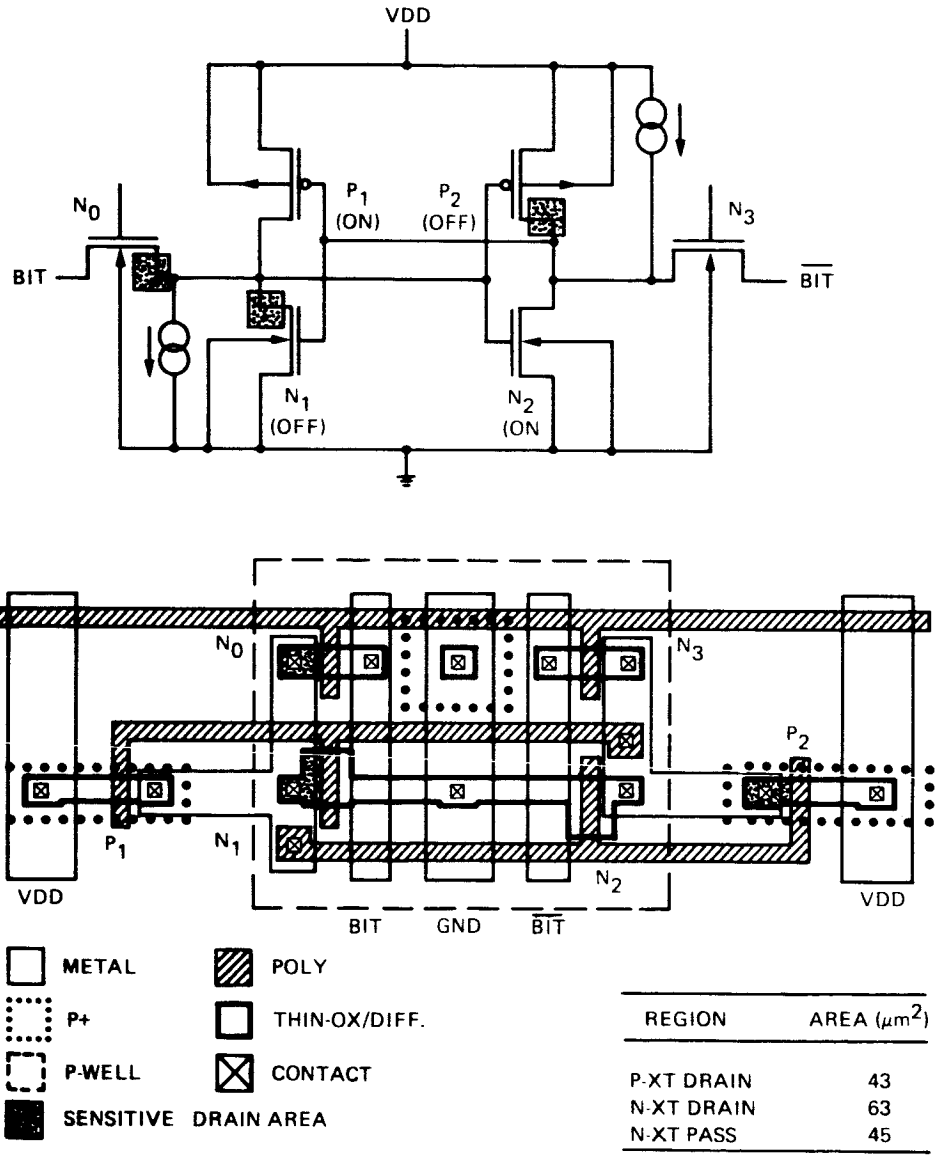


$a \geq 100$ ns
$b \geq 400$ ns
$c \geq 400$ ns
$a \leq d < a + \frac{1}{2}b$

DATA VALID BETWEEN
(a + b) AND (a + b + c)

$e \geq 100$ ns
$f \geq 400$ ns
$g \geq 400$ ns

Figure 2.6-4. CRRES RAM timing diagram



REGION	AREA (μm^2)
P-XT DRAIN	43
N-XT DRAIN	63
N-XT PASS	45

Figure 2.6-5. CRRES RAM memory cell layout and equivalent circuit

	(VDD)				
	5V	4V	3V	2V	1V
Q_C (pc)	0.8	0.7	0.5	0.3	0.2
R (UPSETS/MO/1024 CELLS)	0.3	-	1	-	3

N-CHANNEL XT $A_D = 63\,\mu m^2$, $W = 12\,\mu m$, $L = 3\,\mu m$

P-CHANNEL XT $A_D = 43\,\mu m^2$, $W = 4.5\,\mu m$, $L = 3\,\mu m$

Figure 2.6-6. SEU predictions

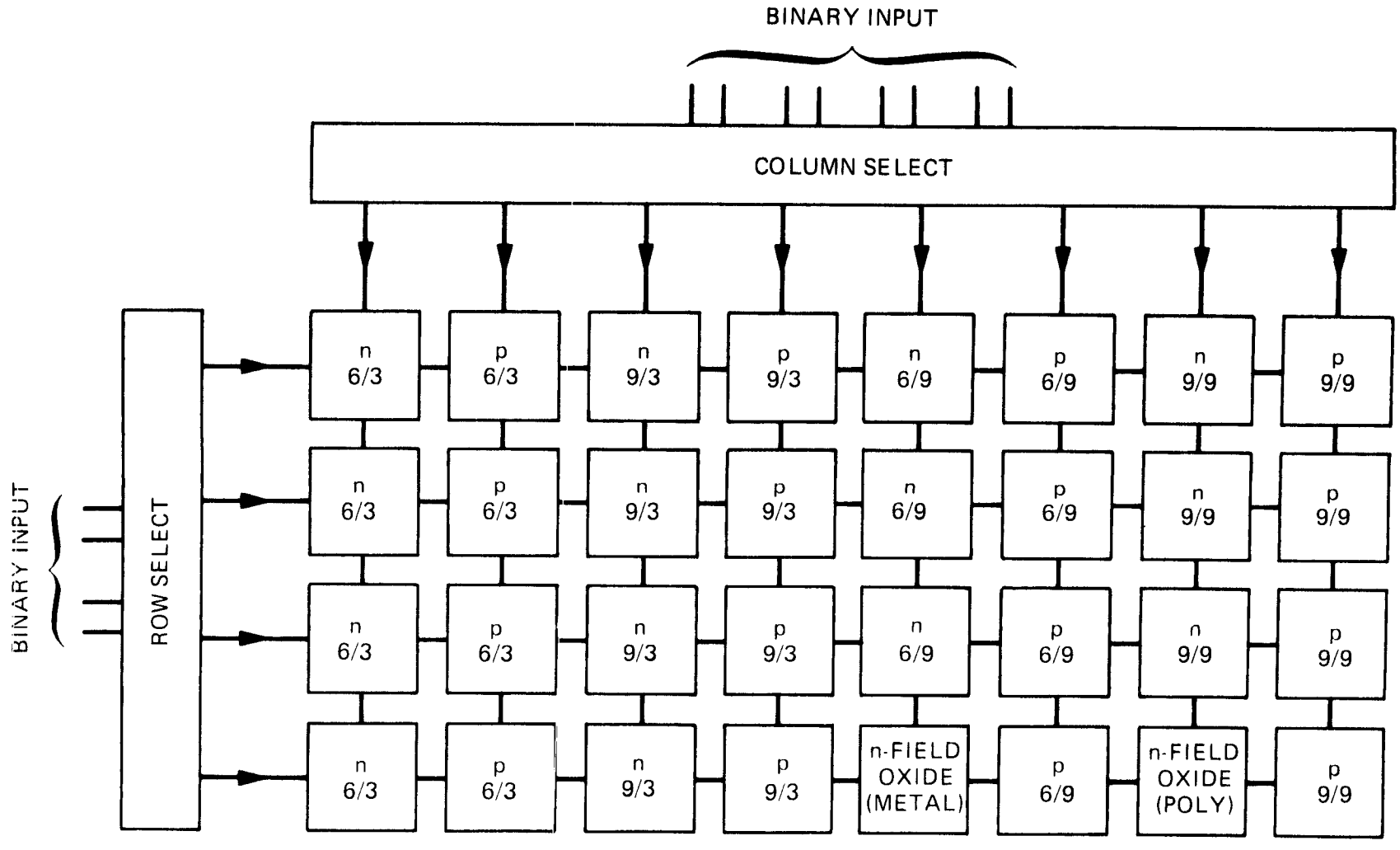


Figure 2.6-7. Transistor matrix (transistor-level description)

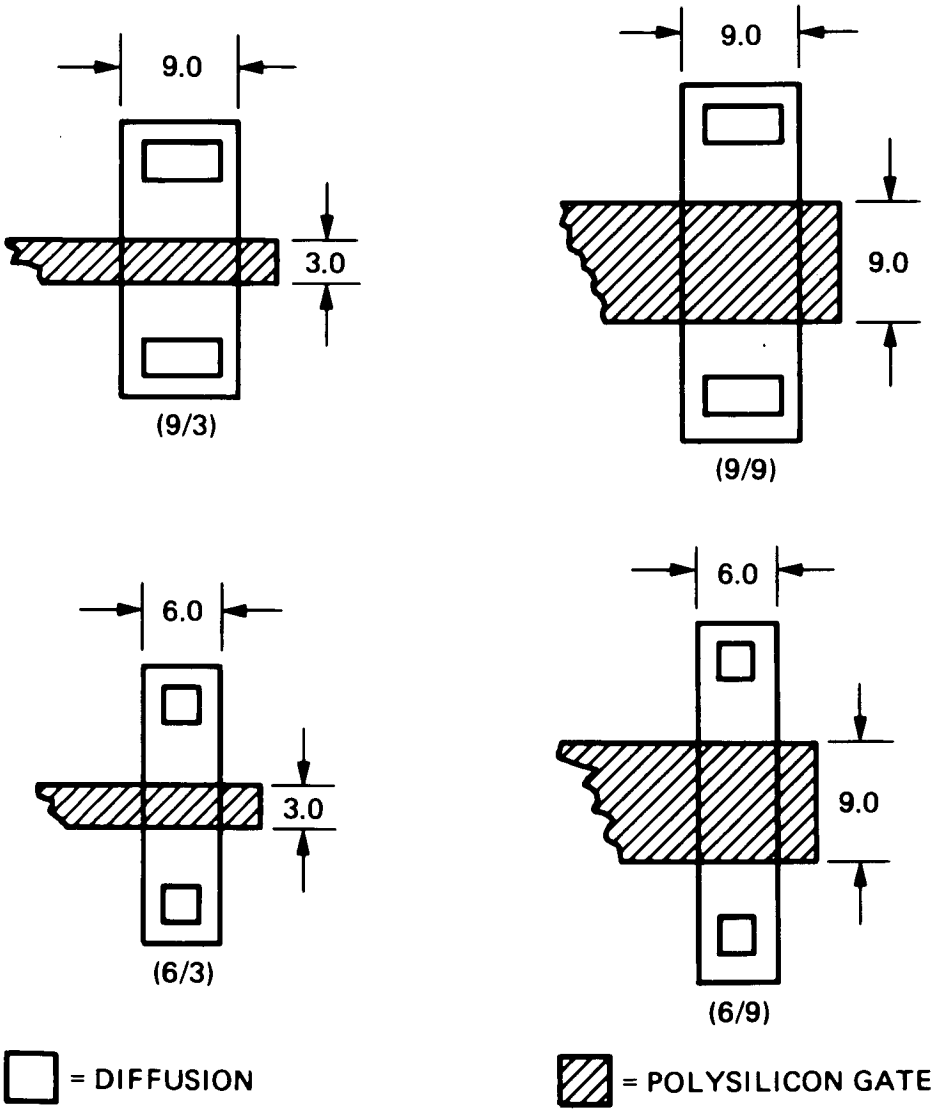


Figure 2.6-8. Geometry of cells in transistor matrix

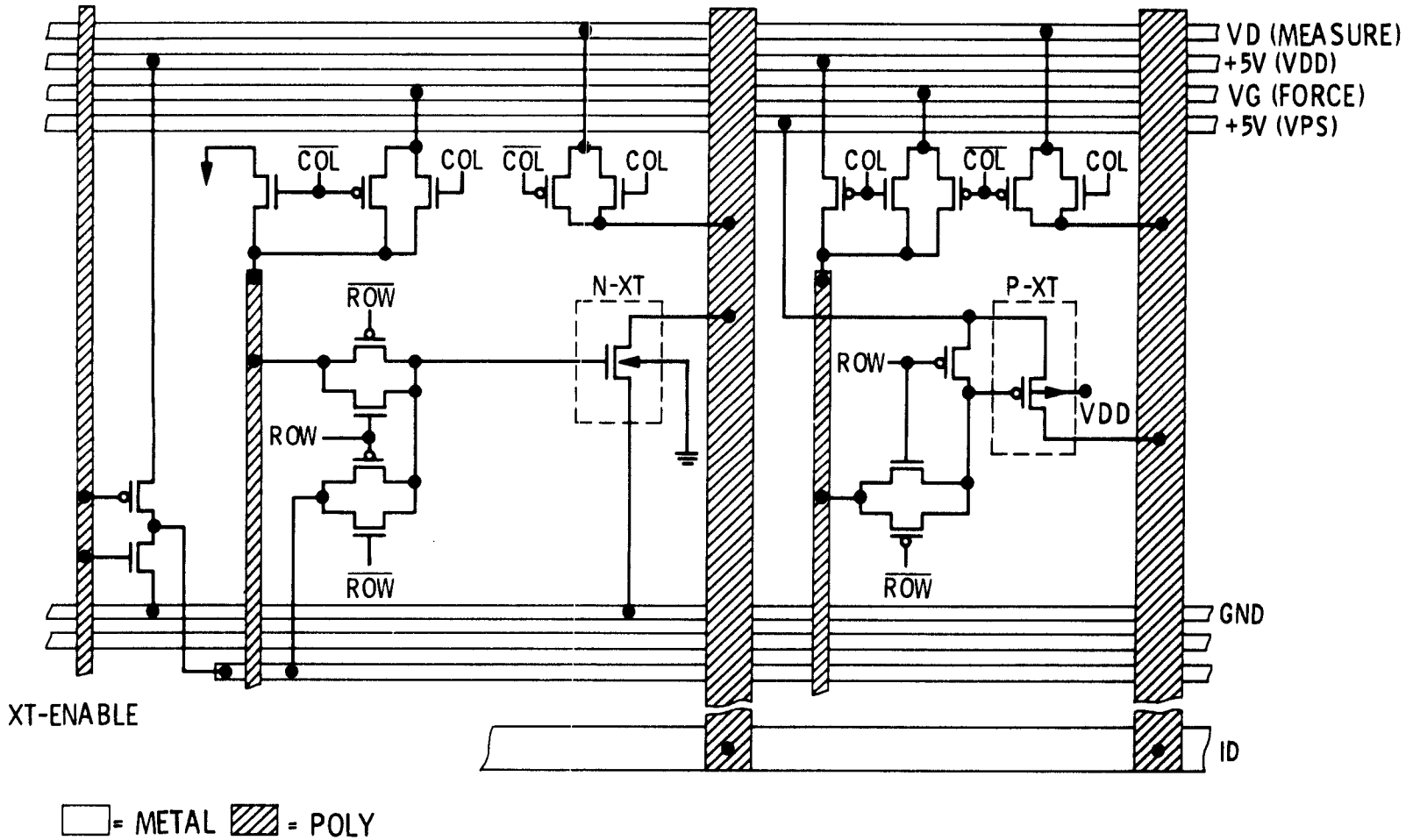


Figure 2.6-9. Transistor placement in XT-matrix

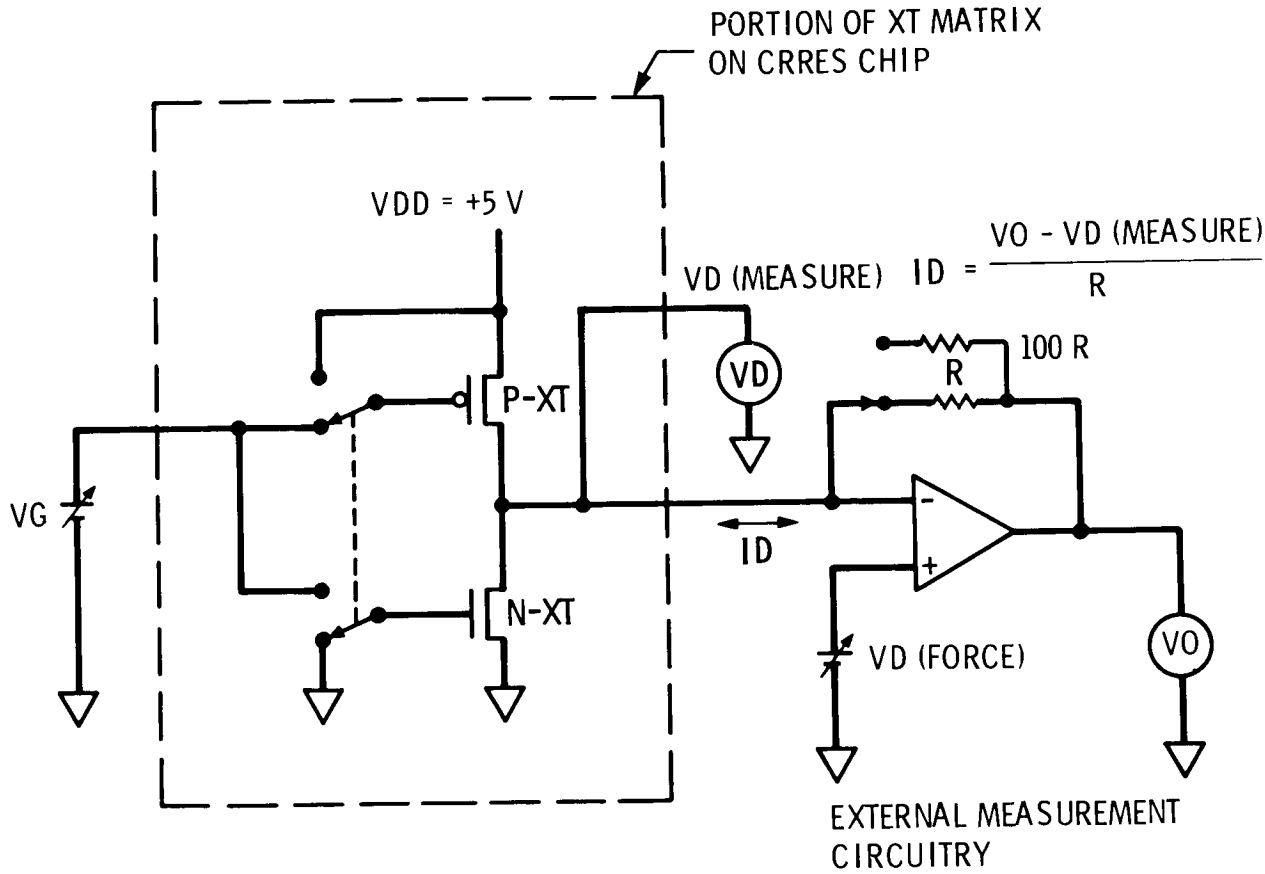


Figure 2.6-10. Measurement configuration for the addressable transistor matrix

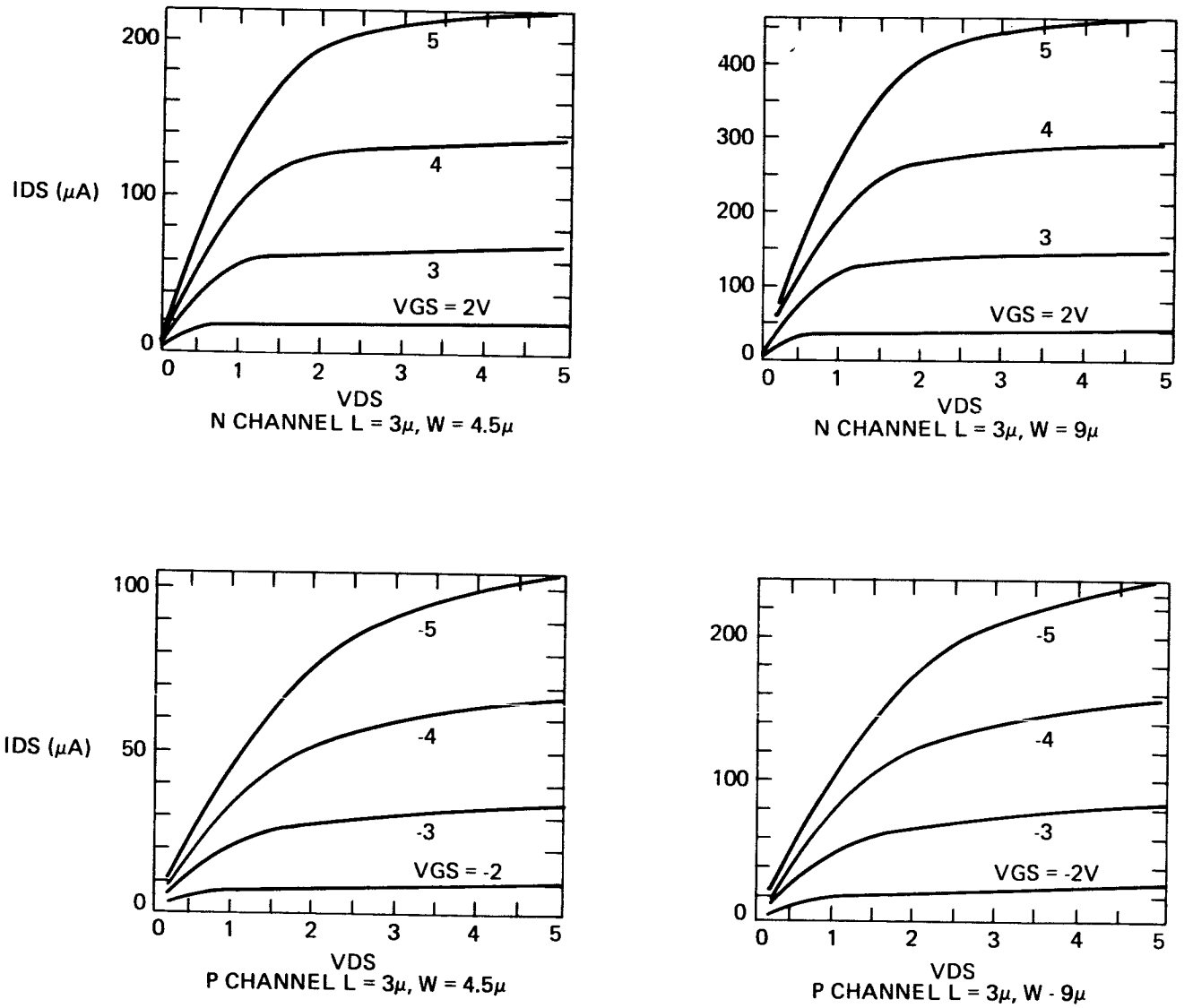


Figure 2.6-11. Transistor matrix test results

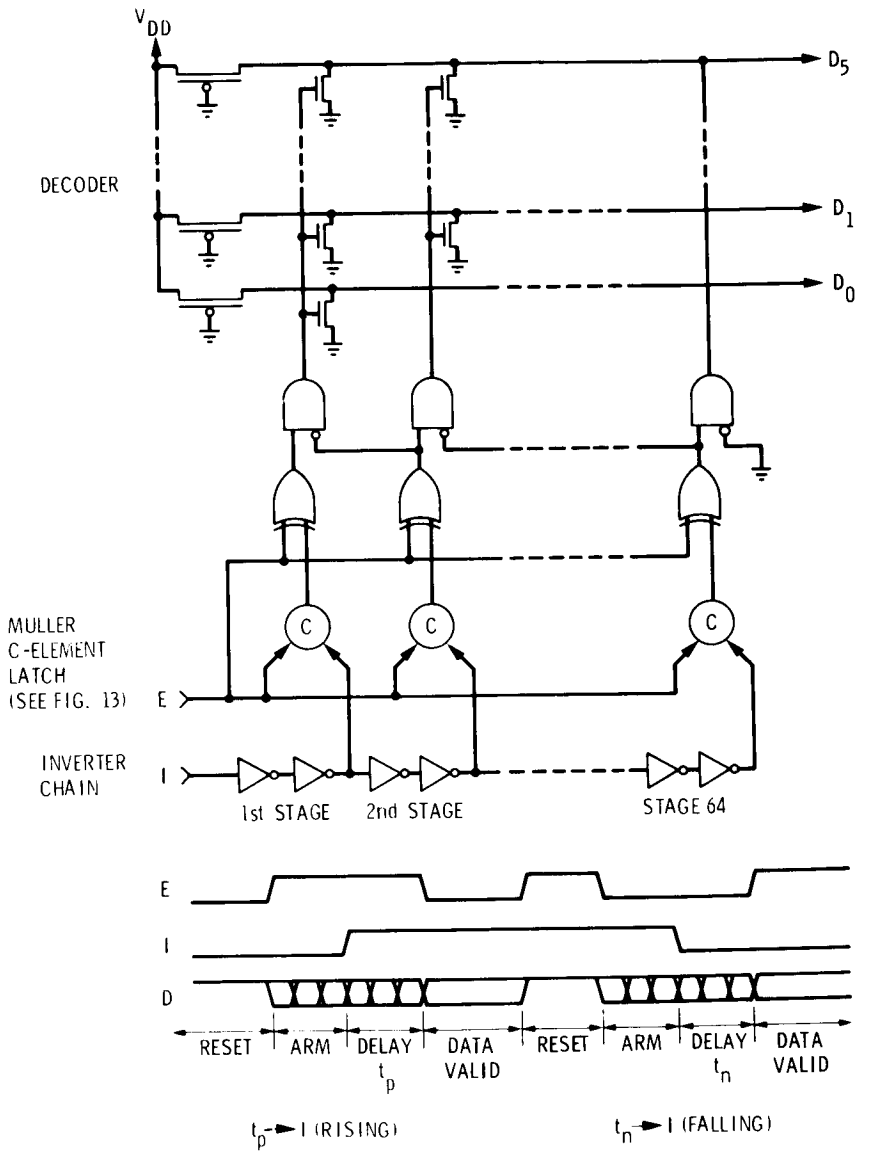
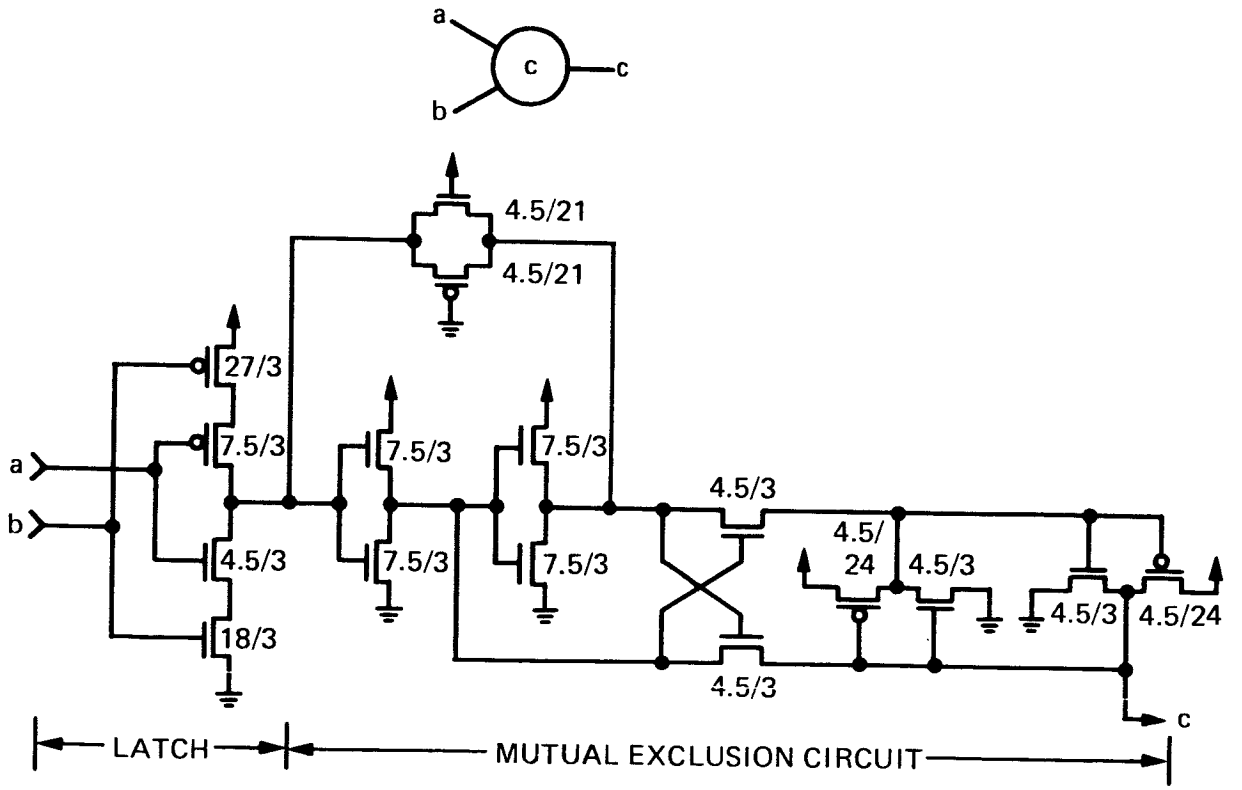


Figure 2.6-12. Timing sampler



W/L = GATE WIDTH/GATE LENGTH EXPRESSED IN MICROMETERS

Figure 2.6-13. Muller C element latch

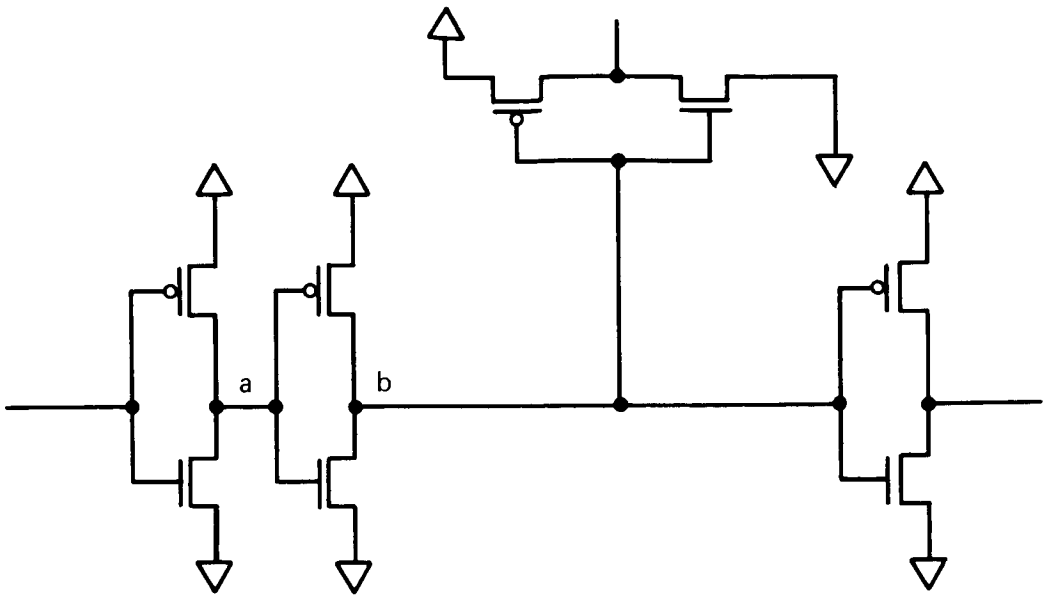


Figure 2.6-14. Loaded inverter pair

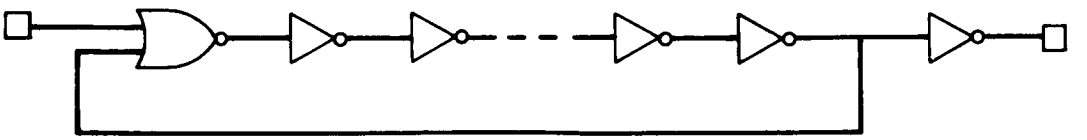


Figure 2.6-15. Ring oscillator

Table 2.6-1. CMOS Timing Sampler and Ring Oscillator Results

CHIP	LOT	TIMING SAMPLER			RING OSCILLATOR		
		T_{pd}(1)	tau-n (eq. 3)	tau-P (eq. 4)	T_{pd}(2)	tau-n (eq. 7)	tau-P (eq. 8)
1	A	2.50	0.313	0.521	2.47	0.307	0.512
2	A	3.14	0.392	0.654	3.07	0.382	0.637
3	A	2.55	0.318	0.530	2.41	0.300	0.500
4	A	2.66	0.332	0.553	3.15	0.393	0.654
5	B	3.40	0.426	0.709	3.35	0.418	0.696
6	B	3.15	0.394	0.656	3.07	0.382	0.637
7	B	3.48	0.435	0.725	3.32	0.414	0.690
8	B	3.54	0.442	0.737	3.27	0.408	0.679
9	B	3.37	0.422	0.703	3.28	0.411	0.685

NOTES: (1) RISING AND FALLING EDGE DELAYS ARE THE SAME:
$T_{pd} = \tau pd\_ip- = \tau pd\_ip+$

(2) THE ON AND OFF TIMES ARE THE SAME:
$T_{pd} = \tau(ON) = \tau(OFF)$

(3) ALL DELAY UNITS ARE IN NANOSECONDS

Table 2.6-2. CMOS CRRES Chip Fabrication Summary

DESIGN	RUN #	SUBMITTED	RECV'D	RAM	MATRIX	T.S.	COMMENTS
Crreschip	M39D	9/5/83	12/19/83	DE	N/A	N/A	RAM only fab.
	M39H	*	2/3/84	DE	N/A	N/A	RAM only fab.
	M3BM	*	3/7/84	DE	N/A	N/A	RAM only fab.
Crreschp1	M3C0	11/30/83	2/6/84	DE	N/A	N/A	RAM only fab.
	M41V	*	5/30/84	DE	N/A	N/A	RAM only fab.
Crreschp2	M41V	1/16/84	5/30/84	7/12	N/A	N/A	Mask & fab. prob.
Crreschp7	M44E	3/30/84	6/20/84	3/12	DE	DE	Fab. problems
Crreschp8	M4GM-1	4/26/84	9/10/84	0/8	DE	DE	Mask & fab. prob.
	M4GM-2	*	9/10/84	2/8	DE	DE	Mask & fab. prob.
Creschp10	M4GM-1	5/23/84	9/10/84	0/2	DE	DE	Mask & fab. prob.
	M4GM-2	*	9/10/84	1/2	DE	DE	Mask & fab. prob.
Creschp11	M4GM-1	6/19/84	9/10/84	3/7	DE	5/7	Mask & fab. prob.
	M4GM-2	*	9/10/84	4/7	DE	3/7	Mask & fab. prob.

* = additional fabrication run submitted by MOSIS

DE = design error prevented circuit from being tested

N/A = not applicable because circuit not included

x/x = operational circuits/chips received

APPENDIX 2.6.A

CRRESCHIP PIN DESIGNATIONS

* - pin unused and tied to ground
\# - pin unused and left unconnected
*** - 4-kbit RAM only

All logic connections are 5 V = true unless otherwise noted

PIN #	NAME	TYPE	DESCRIPTION
1	VDD	POWER (+5)	RAM, XT Matrix select circuit and Pad Ring Vdd
2	XT VPS	POWER (+5)	XT Matrix P-channel XT sources
3	A7 ***	LOGIC IN	4-kbit RAM Address bus, bit 7
4	XT VG	ANALOG IN	XT Matrix gates
5	6TINV OUT	ANALOG OUT (#)	6 Terminal Inverter output
6	6TINV SUB	POWER (+5)	6 Terminal Inverter substrate
7	6TINV VDD	POWER (+5)	6 Terminal Inverter Vdd
8	6TINV P-GATE	ANALOG IN (*)	6 Terminal Inverter P-XT gate
9	6TINV VSS	POWER (GND)	6 Terminal Inverter Vss
10	6TINV WELL	POWER (GND)	6 Terminal Inverter P-Wells
11	XT WELL	POWER (GND)	XT Matrix P-Wells
12	6TINV N-GATE	ANALOG IN (*)	6 Terminal Inverter N-XT gate
13	Spare	Spare (*)	
14	Spare	Spare (*)	
15	TSI	LOGIC IN	Timing Sampler input
16	TSE	LOGIC IN	Timing Sampler enable
17	RO ENABLE	LOGIC IN	Ring Oscillator enable
18	RO OUT	LOGIC OUT (#)	Ring Oscillator output
19	Spare	Spare (*)	
20	Spare	Spare (*)	
21	Spare	Spare (*)	
22	Spare	Spare (*)	
23	TS D0	LOGIC OUT	Timing Sampler output, bit 0
24	TS D1	LOGIC OUT	Timing Sampler output, bit 1
25	TS D2	LOGIC OUT	Timing Sampler output, bit 2
26	TS D3	LOGIC OUT	Timing Sampler output, bit 3
27	TS D4	LOGIC OUT	Timing Sampler output, bit 4
28	TS D5	LOGIC OUT	Timing Sampler output, bit 5
29	GROUND	POWER (GND)	Chip ground
30	DQ 15	LOGIC IN/OUT	RAM Data bus, bit 15
31	DQ 14	LOGIC IN/OUT	RAM Data bus, bit 14
32	DQ 13	LOGIC IN/OUT	RAM Data bus, bit 13
33	DQ 12	LOGIC IN/OUT	RAM Data bus, bit
34	DQ 11	LOGIC IN/OUT	RAM Data bus, bit 11
35	DQ 10	LOGIC IN/OUT	RAM Data bus, bit 10
36	DQ 9	LOGIC IN/OUT	RAM Data bus, bit 9

37	DQ 8	LOGIC IN/OUT	RAM Data bus, bit 8
38	DQ 7	LOGIC IN/OUT	RAM Data bus, bit 7
39	DQ 6	LOGIC IN/OUT	RAM Data bus, bit 6
40	DQ 5	LOGIC IN/OUT	RAM Data bus, bit 5
42	DQ 3	LOGIC IN/OUT	RAM Data bus, bit 3
43	DQ 2	LOGIC IN/OUT	RAM Data bus, bit 2
44	DQ 1	LOGIC IN/OUT	RAM Data bus, bit 1
45	DQ 0	LOGIC IN/OUT	RAM Data bus, bit 0
46	S	LOGIC IN	RAM chip select
47	W	LOGIC IN	RAM read/write select
48	E	LOGIC IN	RAM read/write enable
49	EP	LOGIC IN	RAM precharge control (0 = precharge on)
50	A5	LOGIC IN	RAM Address bus, bit 5
51	A4	LOGIC IN	RAM Address bus, bit 4
52	A3	LOGIC IN	RAM Address bus, bit 3
53	A2	LOGIC IN	RAM Address bus, bit 2
54	A1	LOGIC IN	RAM Address bus, bit 1
55	A0	LOGIC IN	RAM Address bus, bit 0
56	XT ENABLE	LOGIC IN	XT Matrix select
57	A6 ***	LOGIC IN	4-kbit RAM Address bus, bit 6
58	R1	LOGIC IN	XT Matrix row address, bit 1
59	R0	LOGIC IN	XT Matrix row address, bit 0
60	C0	LOGIC IN	XT Matrix col. address, bit 0
61	C1	LOGIC IN	XT Matrix col. address, bit 1
62	C2	LOGIC IN	XT Matrix col. address, bit 2
63	XT ID	ANALOG IN	XT Matrix XT drain current
64	XT VD	ANALOG OUT	XT Matrix XT drain voltage

APPENDIX 2.6.B

INTRINSIC GATE DELAY (TAU) MODEL

In this appendix we derive the equation for the intrinsic delay of an inverter that drives a capacitive load. The time delay is related to the drive capability of the inverter and hence will reflect whether the inverter is pulling its output from a low to a high state or vice versa.

For a rising step pulse input to the inverter, shown in Figure 2.6.B-1, the output voltage can be found by considering that immediately after the pulse the pull-up transistor Qp is off and the pull-down transistor Qn is on.

In this example, the drain of Qn is initially biased at Vdd so that Qn is in saturation (i.e., VDn >> VDsat = VGn - VTn). For this n-channel transistor (subscript n), the drain current is:

$$IDn = (UnCoWn)(VGn - VTn)^2/(2Ln) \qquad (1)$$

where VGn = Vi = Vdd, U is the channel mobility, W is the channel width, L is the channel length, VT is the threshold voltage, and Co is the gate oxide capacitance per unit area. Since Qp is off, the current through Qn is equal to the current drawn from the load capacitance, Cz, or

$$Iz = Cz \cdot dVo/dt \qquad (2)$$

The load capacitance is equal to the input capacitance of subsequent stages, the junction capacitance of the drains of Qn and Qp, and the wire capacitance [1]. The input capacitance of the subsequent stages is given by the fanout times the sum of the gate capacitances of both the pull-up and pull-down transistors.

By combining Eqs. (1) and (2), using the Kirchhoff current law, integrating with respect to time, and evaluating the constant of integration at Vo(t = 0) = Vdd, the output voltage of the inverter while Qn remains in saturation is

$$Vo = Vdd - [(UnCn)/(2Ln^2Cz)](Vdd - VTn)^2t \qquad (3)$$

where Cn = CoWnLn is the gate capacitance of Qn. According to Eq. (3), the time t1 for the inverter output to decay to zero is:

$$t1 = (Cz/Cn)TAUn \qquad (4a)$$

where the intrinsic delay for Qn is defined as:

$$TAUn = 2VddLn^2/[Un(Vdd - VTn)^2] \qquad (4b)$$

Thus TAUn is defined for the case where Qn is assumed to remain in saturation as the output of the inverter goes to zero. As seen in Figure 2.6.B-2, once

the inverter output is no longer in saturation, the output decays asymptotically to zero.

Using similar reasoning for a falling step pulse input to the inverter, the intrinsic delay for Qp is

$$TAUp = 2Vdd\ Lp^2/[Up(Vdd - |VTp|)^2] \tag{5}$$

where Qn is off and Qp must pull the inverter output high.

The TAUn and TAUp are used to estimate the propagation delay for inverter operation. For a falling inverter output the propagation delay is estimated as

$$TAUpd- = (Cz/Cn)TAUn \tag{6}$$

and for a rising inverter output the propagation delay is estimated as

$$TAUpd+ = (Cz/Cp)TAUp \tag{7}$$

where Cp = CoWpLp is the gate capacitance of Qp. It is noted that TAUpd- = t1 derived in Eq. (4a). The utility of Eqs. (6) and (7) is found in that the propagation delay for the inverter can be calculated by simply multiplying the intrinsic gate delay (TAUn or TAUp) by the ratio of the load capacitance (Cz) to the driving transistor capacitance (Cn or Cp).

2.6.B.2 REFERENCE

1. S. M. Kang, "A Design of CMOS Polycells for LSI Circuits," IEEE Trans. on Circuits and Systems, CAS-28, 838-843 (1981).

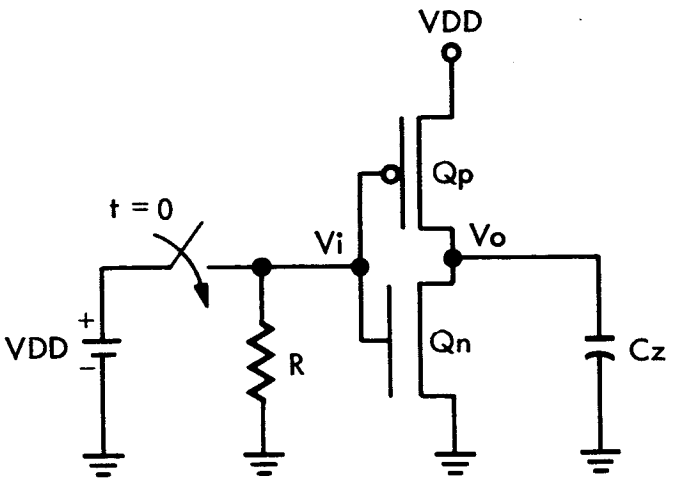


Figure 2.6.B-1. CMOS inverter used to calculate the intrinsic gate delay

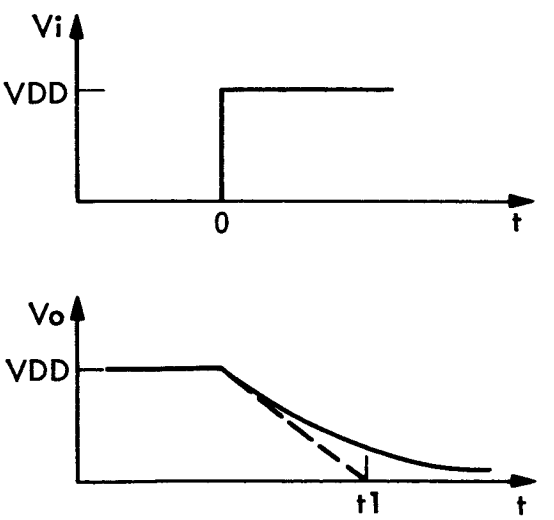


Figure 2.6.B-2. The input and output signal for a CMOS inverter

SECTION 3

DISCUSSION AND PLANS

3.1 OVERVIEW

This effort has developed a number of test structures and the associated test methodology. In some cases, the methods are mature and ready for routine use (e.g., the Pinhole Array Capacitor). In other cases, the methods are developmental and require additional data to establish the validity of the method (e.g., the time-dependent dielectric breakdown). In still other cases, the methods have not been abstracted from the literature and reduced to practice (e.g., subthreshold leakage current, and transistor capacitance). In the follow-on effort we will be working to fill in the gaps in the measurement technology with an eye to establishing a comprehensive test strip and/or test structure methodology that can be used in routine use.

One of the changes that appear mandatory is the change from the currently used 2-by-N probe pad configuration [1] shown in Figure 3-1. This configuration has served well for over five years. It was originally developed when test chip drop-ins were an accepted practice. But with the advent of direct-step-on wafer photolithography (as opposed to whole wafer photolithography), the need to conserve space became paramount. Several GaAs manufacturers have adopted the configuration shown at the bottom of Figure 3-1, and it is seen that there is a 60 percent space savings. The disadvantage of the more compact structure is that bonding is now impossible, whereas with older structures, bonding was possible. We intend to convert our test structures to the compact probe-pad configuration shown in Figure 3-1.

The major complaint concerning the use of test structures is the lack of comparison between test structure results and circuit performance. Test chip and test strips [2] are in routine use in the industry, but there is little data that confirms the comparison. For example, Bernard [3] shows how ring oscillator performance compares with contact resistance and Ipri [4] shows how inverter chains can reveal layout rule limits. To further develop the data base for comparing test structure results and circuit performance, four types of test chips are being arranged on the wafer as shown in Figure 3-2. Two of the chips are CRRES chips. These chips have a timing sampler, RAM, and transistor matrix. The transistor count for the CRRES (RAM1) chip is about 12,000, and for the CRRES (RAM2) chip is about 30,000. The performance of the circuits on these chips will provide the circuit performance data. The test structures are contained on the test chip (CM5041) and contain structures similar to those described in Section 2.4.2. A transistor matrix has been included to allow a detailed parameter extraction of the DC transistor parameters after irradiation.

The overall approach to developing the wafer acceptance procedures is illustrated in Figure 3-3. It indicates that approximately seven dedicated 3-μm CMOS runs are required to establish the data base in support of the case studies in which the correlation between test structure results and circuit performance will be established. In this manner we will be able to develop wafer acceptance procedures that are based on an engineering data base.

3.2 REFERENCES

1. M. G. Buehler, "Comprehensive Test Patterns With Modular Test Structures: The 2 by N Probe-Pad Array Approach," Solid State Technol., 22, 89-94 (October 1979).

2. C. Alcorn, D. Dworak, N. Haddad, W. Hendley, and P. Nixon, "Kerf Test Structure Designs for Process and Device Characterization," Solid State Technol., 28, 229-235 (May 1985).

3. J. Bernard, "The IC Yield Problem: A Tentative Analysis for MOS/SOS Circuits," IEEE Trans. Electron Devices, ED-25, 939-944 (1978).

4. A. C. Ipri and J. C. Sarace, "Integrated Circuit Process and Design Rule Evaluation Techniques," RCA Rev., 38, 323-350 (1977).

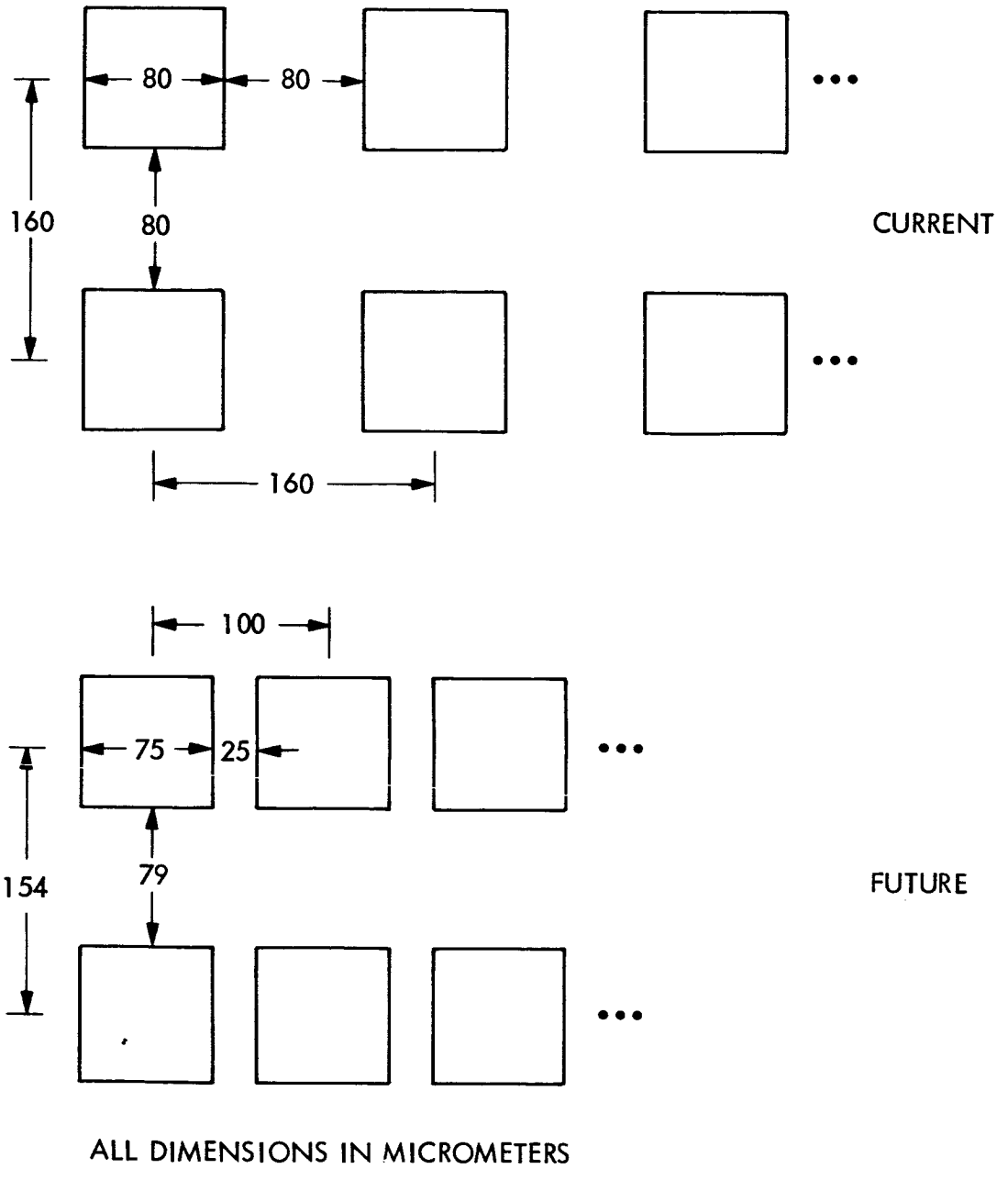


Figure 3-1. Current and future 2-by-N probe pad configurations

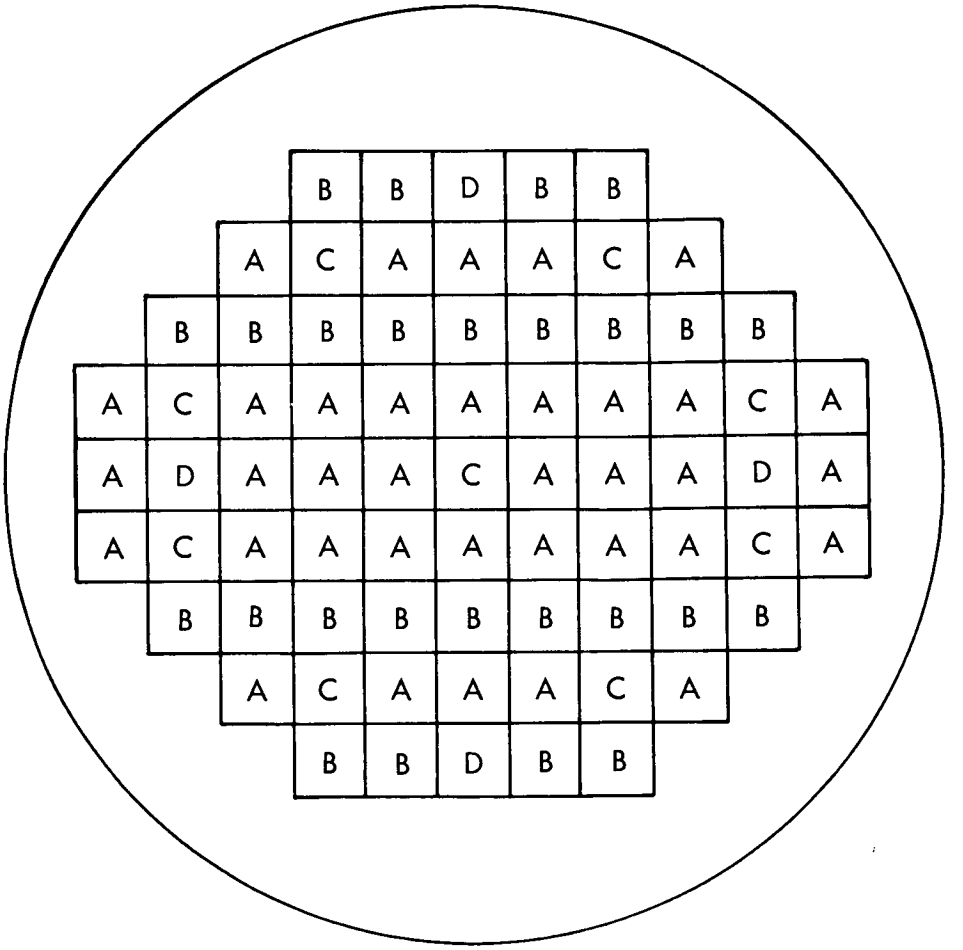


Figure 3-2. Four types of chips arranged on a 4-inch-diameter silicon wafer that will be used to establish a correlation between test structure results and integrated circuit performance

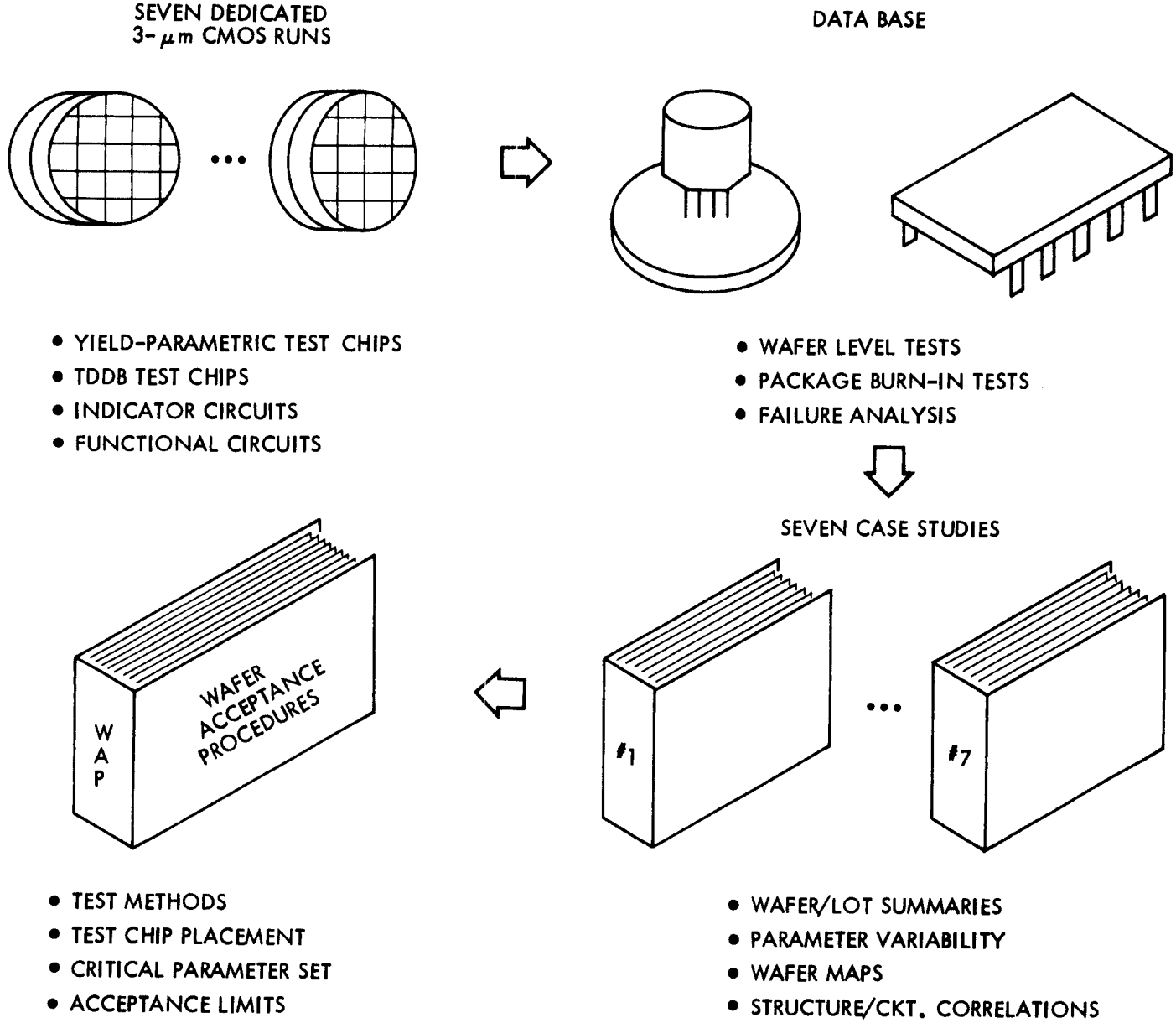


Figure 3-3. Suggested procedure to establish a data base for establishing wafer acceptance procedures

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No. 85-76	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle
Product Assurance Technology for Custom LSI/VLSI Electronics
Report for Period: October 1982 - September 1984

5. Report Date
June 1985

6. Performing Organization Code

7. Author(s) M.G. Buehler, B.R. Blaes, G.A. Jennings, B.T. Moore, R.H. Nixon, C.A. Pina, H.R. Sayah, *

8. Performing Organization Report No.

9. Performing Organization Name and Address
JET PROPULSION LABORATORY
California Institute of Technology
4800 Oak Grove Drive
Pasadena, California 91109

10. Work Unit No.

11. Contract or Grant No.
NAS7-918

13. Type of Report and Period Covered
External Report
JPL Publication

12. Sponsoring Agency Name and Address
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
Washington, D.C. 20546

14. Sponsoring Agency Code
RE191 BW32351200700

15. Supplementary Notes
* M.W. Sievers, and N.F. Stahlberg

16. Abstract The work described in this report represents the collaborative efforts of integrated-circuit (IC) parts specialists, device physicists, test-chip engineers, and fault-tolerant-circuit designers. Their efforts were focused on developing the technology for obtaining custom ICs from CMOS-bulk silicon foundries using a universal set of layout rules. In pursuit of this goal the technical efforts were guided by the requirement to develop a 3-μm CMOS test chip for the Combined Release and Radiation Effects Satellite (CRRES). This chip contains both analog and digital circuits for characterizing the space-radiation-induced shifts in CMOS transistor parameters and inverter propagation delays, and for characterizing the single-event-upset (SEU) rates of static random access memories (SRMs). The development employed all the elements required to obtain custom circuits from silicon foundries, including circuit design, foundry interfacing, circuit test, and circuit qualification. The technical accomplishments of this effort include: (a) A critical review of the military IC qualification standards. The review indicated that many of these standards are not applicable to the procurement of custom ICs. (b) An assessment of the test time and area required by CMOS-bulk test structures that are to be included in parameter extraction and yield analysis test chips. (c) The development of a MOSFET parameter extraction procedure, called JMOSFIT, which allows the extraction of physically meaningful parameters for a SPICE-like circuit simulator. (d) The fault modeling of pinhole-array-capacitor defects in terms of a gate-to-silicon resistive short that terminates in an n-diffusion pocket in the silicon. (e) The time response analysis of CMOS gates (inverters, NANDs, NORs, and flip-flops) for various resistive shorts which indicates the resistive values needed to cause a faulty response. (f) The development of a CMOS-bulk test chip for CRRES.

17. Key Words (Selected by Author(s))	18. Distribution Statement
Electronics and Electrical Engineering; Methods and Equipment (General); Solid-State Physics	Unclassified; unlimited

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified		